INTERLIBRARY LOAN REQUEST FORM

Borrower's Name	Nancy Johnson	Org. or A.U.	1642	Phone	305-5860
Serial Number	09/017,715	Date of Request	9.14.99	Date Needed By	ASAP

Please Attach Copy of Abstract, Citation, Or Bibliography. If Available Please Provide Complete Citation. Only One Request Per Form.

Author/Editor:	Ji, H
Journal/Book Title:	Cancer Research
Article Title:	
Volume (Issue):	57(4)
Pages:	759-764
Year of Publication:	1997
Publisher:	
Remarks:	applicant's work

STIC Use Only

Accession Number: _____

LIBRARY ACTION	LC		NAL		NIH		NLM		NBS		PTO		OTHER	
	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd	1st	2nd
Local Attempts														
Date														
Initials														
Results														
Examnr. Called														
Page Count														
Money Spent														

Provided By: Source and Date

Ordered From: Source and Date

Remarks/Comments

-1st & 2nd denotes times taken to a library
-FX - Means Faxed to us
-O/N - Under NLM=Overnight Service

LATE-BREAKING RESEARCH SESSION AT THE AACR ANNUAL MEETING

Tuesday, April 13, 1999

Time has been set aside for the presentation of a few definitive reports of highly significant and timely findings in the field. Criteria for the selection of these presentations and instructions for submission of abstracts are as follows:

INSTRUCTIONS FOR SUBMISSION OF LATE-BREAKING ABSTRACTS

1. The work to be presented must be of major novelty and significance, *e.g.*, the characterization of a new gene in familial cancer or the discovery of a new diagnostic marker, and should not have been previously published in a peer-reviewed scientific journal or presented at a national meeting.

2. The abstract must be sponsored by an AACR member in good standing (dues paid for 1999).

3. Each member in good standing may sponsor only **one** abstract for this session whether or not he or she sponsored an abstract last November for the regular annual meeting program. If an associate member is the **sponsor**, the abstract must also be **endorsed** by an active or corresponding member in good standing. In this case, the **endorser** does **not** forfeit the opportunity to **sponsor** a late-breaking abstract.

4. Abstracts must be typed on **one** side of **one** sheet of white paper.

5. All text on the page must fit within an area 6 1/2" wide and 9" high (16.5 cm X 22.9 cm) with margins of at least 1" (2.5 cm) on the top, bottom, and sides of the page.

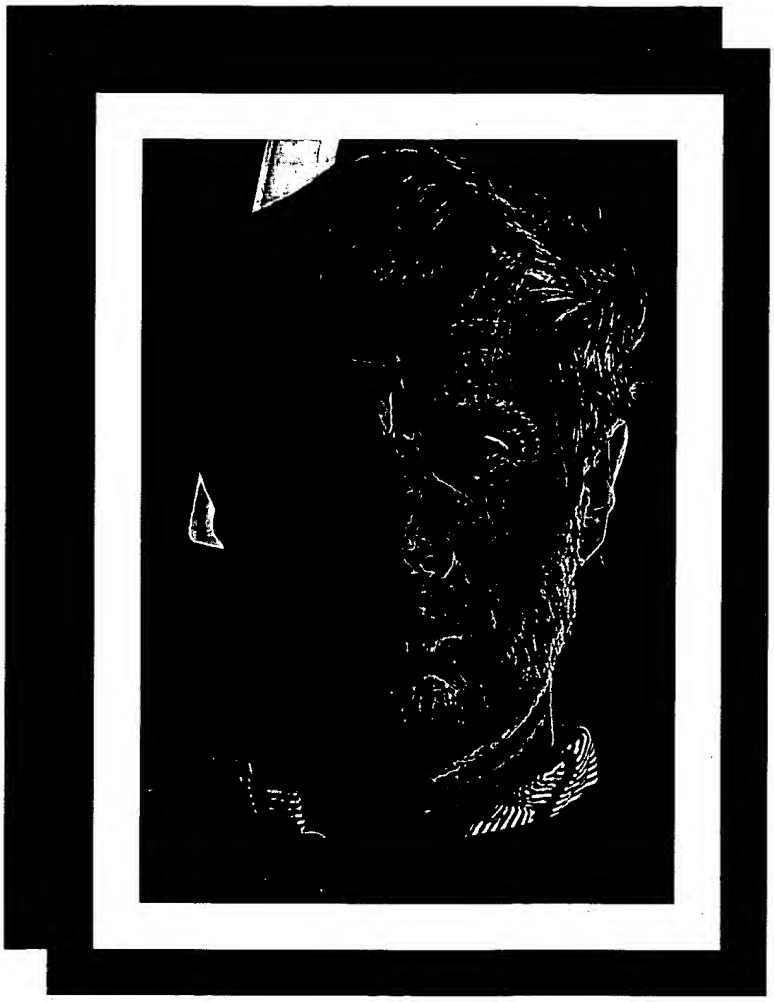
6. Each abstract must be accompanied by a covering letter from the sponsor explaining why the work is novel and significant enough to be considered for this late-breaking research session and certifying that the findings became available **after** the annual meeting abstract deadline of November 6, 1998. This letter must contain the sponsor's complete mailing address, FAX number, and E-mail address (if available) so that we can communicate the scheduling decision of the Program Committee.

7. Abstracts and covering letters must be received in the AACR Office by 5:00 p.m. Eastern Time on **March 12, 1999.** FAX transmissions are **not** acceptable. Carrying envelopes should be clearly marked "Late-Breaking Abstract," and should be addressed to American Association for Cancer Research, Public Ledger Building, Suite 826, 150 South Independence Mall West, Philadelphia, PA 19106-3483. If you wish to receive acknowledgment of receipt of your abstract, enclose a self-addressed post card with appropriate postage affixed. Accepted abstracts will not be published since they will be received after the *Proceedings of the American Association for Cancer Research* has been printed; however, they will be distributed at the session in Philadelphia.

8. A special subcommittee of the Program Committee appointed by President Webster K. Cavenee will select the papers to be presented. Presenters of accepted papers will be notified via FAX no later than **March 29, 1999**.

Cancer Research

AN OFFICIAL JOURNAL OF THE AMERICAN ASSOCIATION FOR CANCER RESEARCH



February 1, 1999
Volume 59 • Number 3
PP. 507-769
ISSN 0008-5472 • CNREA 8

[CANCER RESEARCH 59, 742–747, February 1, 1999]

Stimulation of Breast Cancer Invasion and Metastasis by Synuclein γ[1]

Tongli Jia, Yiliang E. Liu, Jingwen Liu, and Y. Eric Shi[2]

Departments of Pediatrics [T. J., Y. E. L., Y. E. S] and Pathology [Y. E. S], Long Island Jewish Medical Center, The Long Island Campus for the Albert Einstein College of Medicine, New Hyde Park, New York 11040, and Palo Alto VA Health Care System, Palo Alto, California 94304 [J. L.]

ABSTRACT

We recently identified and cloned novel breast cancer-specific gene *BCSG1* by direct differential cDNA sequencing. BCSG1 has a great sequence homology with the Alzheimer's disease-related neural protein synuclein (SNC); thus, it was also named SNC-γ. Overexpression of SNC-γ in breast cancer cells leads to a significant increase in motility and invasiveness *in vitro* and a profound augmentation of metastasis *in vivo*. Our data suggest that this member of the neural protein SNCs might have important functions outside the central nervous system and may play a role in breast cancer progression.

INTRODUCTION

If sufficiently characterized, the identification of quantitative changes in gene expression that occur in the malignant mammary gland may yield novel molecular markers that may be useful in understanding breast cancer development and progression (1). Within this context, we have previously reported the isolation of differentially expressed genes in cDNA libraries from normal breast tissue and infiltrating breast cancer using the expressed sequence tag-based differential cDNA sequencing approach (2, 3). Of the many putative differentially expressed genes (2, 3), *BCSG1*, which was identified as a group of expressed sequence tags specifically expressed in the mammary gland relative to other organs and abundantly expressed in a breast cancer cDNA library but scarcely seen in a normal breast cDNA library, was identified as a putative breast cancer-specific gene (2).

Interestingly, BCSG1 revealed no homology to any other known growth factors or oncogenes; however, BCSG1 revealed extensive sequence homology to the AD[3]-related neural proteins called SNCs that are expressed mainly in the brain and localized to presynaptic terminals (4–7). The pathological hallmark of AD is amyloid deposition in neurotic plaques and blood vessels (8). Two major intrinsic constituents of amyloid are a 39–43-amino acid peptide named the Aβ component (8) and the recently identified non-Aβ component (4). The non-Aβ component of the AD precursor was cloned from a human brain library (4) and named SNCA because it shares a 95% sequence homology with rat SNC. Recently, a second SNC named SNCB was cloned from human brain, and it has a 61% sequence identity with SNCA (6). The previously identified BCSG1, which is also highly expressed in the brain (2), has a 54 and 56% sequence identity with SNCA and SNCB, respectively, and has been renamed SNCG (9). Thus, the previously unrecognized homology between these proteins defines a family of human brain SNCs that currently has three members. Although SNCs are abundant proteins expressed in presynaptic terminals and are strongly associated with amyloid plaque in AD and Lewy body in PD (10), their functions have not yet been defined. SNCA aggregation may be important in the etiology and pathogenesis of neurodegenerative disorders such as AD and PD (10). During its identification as a breast cancer-specific gene, we previously demonstrated stage-specific SNCG expression as follows: (*a*) SNCG was undetectable in normal or benign breast lesions; (*b*) SNCG showed partial expression in ductal carcinoma *in situ*; and (*c*) SNCG was expressed at an extremely high level in advanced infiltrating breast cancer. The effects of SNCG on breast cancer growth and metastasis were investigated in the current studies.

MATERIALS AND METHODS

Transfection. Full-length SNCG cDNA was inserted into a pCI-neo mammalian expression vector, and the resulting vector was transfected into MDA-MB-435 cells as described previously (3, 11).

Preparation of CM. All of the clones were maintained in subconfluent monolayers with 10% FCS. The medium was discarded, and the monolayers were washed twice with PBS. The monolayers were cultured in the absence of serum in DMEM supplemented with transferrin (1 mg/liter), fibronectin (1 mg/liter), and trace elements (Biofluids, Rockville, MD). After 24 h, the serum-free medium was discarded, and the cells were replenished with the fresh serum-free medium. The CM were collected 30 h later. Media were then centrifuged at 1,200 × *g*, and the supernatants were saved and concentrated approximately 5-fold using an Amicon hollow fiber concentrator with a M_r 10,000 cutoff at 4°C. The protein concentrations of CM were determined and normalized.

MMP Activity. The MMP enzymatic activity of the CM was assayed using a quenched fluorescent substrate Mca-Pro-Leu-Gly-Leu-Dpa-Ala-Arg-NH_2 (Bachem) as described previously (12). The CM were pretreated with APMA for activation (13).

***In Vitro* Invasion and Motility Assay.** As described previously (11), cell invasion and motility were analyzed in a modified Boyden chamber assay using 8-μm polycarbonate membranes coated with 4 mg/ml growth factor-reduced Matrigel.

Tumor Growth in Athymic Nude Mice. A tumorigenic assay was performed in nude mice as described previously (3, 11). Briefly, approximately 0.4×10^6 cells (0.15 ml) were injected into a 5–6-week old female athymic nude mouse (Frederick Cancer Research and Development Center, Frederick, MD). Each animal received two injections, one on each side, in the mammary fat pads between the first and second nipples. Tumor size was determined at weekly intervals by three-dimensional measurements (in millimeters) using a caliper. Only measurable tumors were used to calculate the mean tumor volume for each tumor cell clone at each time point. Animals were sacrificed 32–40 days after injection, when the largest tumors reached about 15 mm in diameter.

Assessment of Regional Lymph Node and Lung Metastasis. As described previously (11), the axillary lymph nodes and lungs of sacrificed animals were excised, weighed, fixed in formalin, embedded in paraffin, and stained with H&E for a microscopic examination for morphological evidence of tumor metastasis. Sections were reviewed and scored by two pathologists.

Antibody Production. The purified synthetic SNCG peptide corresponding to amino acids 101–117 (2) was conjugated and injected into New Zealand rabbits as reported previously (12). The antiserum was purified using SNCG peptide affinity chromatography.

Received 8/24/98; accepted 12/1/98.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by Grant DAMD17-98-8118 from the United States Army Breast Cancer Research Program, Grant RPG-99-028-01-CCE from the American Cancer Society, and Helen and Irving Schneider.

[2] To whom requests for reprints should be addressed, at Department of Pediatrics, Schneider Children Hospital, Long Island Jewish Medical Center, New Hyde Park, NY 11040. Phone: (718) 470-3086; Fax: (718) 470-6744; E-mail: shi@lij.edu.

[3] The abbreviations used are: AD, Alzheimer's disease; APMA, *p*-aminophenylmercuric acetate; MAP, microtubule-associated protein; MMP, matrix metalloproteinase; PD, Parkinson's disease; SNC, synuclein; SNCA, SNC-α; SNCB, SNC-β; SNCG, SNC-γ; CM, conditioned media; OM, oncostatin M.

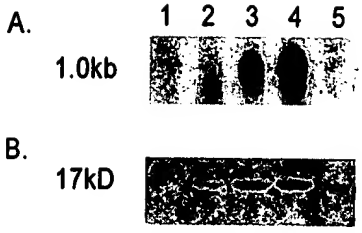


Fig. 1. Transfection of SNCG to MDA-MB-435 cells. *A*, Northern blot. Each lane contained 30 μg of total RNA. *B*, Western blot with an affinity-purified specific SNCG peptide polyclonal antibody. Each lane contained 20 μg of protein. *Lane 1*, neo-435-1; *Lane 2*, SNCG-435-2; *Lane 3*, SNCG-435-1; *Lane 4*, SNCG-435-3; *Lane 5*, neo-435-2.

RESULTS AND DISCUSSION

Transfection of SNCG into MDA-MB-435 Human Breast Cancer Cells. To determine the effects of SNCG on invasion/metastasis, we selected MDA-MB-435 human breast cancer cells as recipients for SNCG-mediated gene transfection due to their lack of detectable SNCG transcript (2) and their highly tumorigenic and aggressive phenotype in nude mice (11). Cells were transfected with a plasmid vector containing a neomycin resistance gene (neo clones) or with the same vector containing full-length SNCG cDNA (SNCG clones). MDA-MB-435 clones expressing SNCG were designated as SNCG-435 clones, and the control neo-transfected cells were designated as neo-435 clones. Fig. 1 shows the Northern blot and Western blot analyses of SNCG expression in selected clones. All selected SNCG-435 clones expressed SNCG mRNA transcripts and proteins. In contrast, none of the neo-435 clones produced any detectable SNCG transcripts and proteins. No changes in morphology were observed in these clones. Based on the level of SNCG expression, we selected SNCG-435-1, SNCG-435-3, neo-435-1, and neo-435-2 clones for the subsequent studies.

***In Vitro* Growth of SNCG-transfected Cells.** To determine whether SNCG overexpression affects the growth of MDA-MB-435 cells, cells from exponentially growing cultures of different MDA-MB-435 clones were seeded in triplicate at 3000 cells/well (24-well plate) in 1 ml of DMEM-5% serum. The growth rates of SNCG-positive SNCG-435-1 and SNCG-435-3 cells were compared with those of SNCG-negative neo-435-1 and neo-435-2 cells in a monolayer culture. No significant differences in growth rate were observed among SNCG-positive and SNCG-negative cells (data not shown).

Metastasis in the Orthotopic Nude Mice Model. Because SNCG was highly expressed in the infiltrating breast cancer cells relative to benign or noninvasive *in situ* carcinomas (2), we were interested in studying whether SNCG is an instigator of metastasis or merely a correlative product during breast cancer progression. The effect of SNCG expression on metastasis was assayed in an *in vivo* orthotopic (mammary fat pad) nude mouse model. Two independent experiments were done to confirm reproducibility, and the data from these experiments are summarized in Table 1. After a lag phase of 10 days, mice given implants of both SNCG-positive and SNCG-negative cells developed tumors. There was no difference in tumor incidence between neo-435 and SNCG-435 clones. Starting at about 20 days after inoculation, tumor necrosis was observed in tumors derived from SNCG-435-1 and SNCG-435-3 cells. Neo-435-1 and neo-435-2 cells also developed some tumor necrosis, but with less intensity. Consistent with the similar *in vitro* growth rates, there was no significant difference in primary tumor size between the neo-435 and SNCG-435 clones at 40 days after injection.

To study tumor dissemination, axillary lymph nodes and lungs were examined physically at autopsy and then subjected to microscopic examination for morphological evidence of tumor cells by light microscopy on H&E-stained paraffin sections. For the axillary lymph node, the average weight was 15 mg for neo-435 mice and 44 mg for SNCG-435 mice. The increased lymph node weight reflects the invaded breast tumors. Representative H&E-stained sections for neo-435 and SNCG-435 lymph nodes are shown in Fig. 2, *A* and *B*. Microscopic examination indicated that SNCG-435-1 and SNCG-435-3 mice showed a significantly higher average lymph node positivity (64 and 77%) compared to that (27%) of SNCG-negative neo-435-1 and neo-435-2 mice (Table 1). For lung metastases, the numbers of visible nodules on the surface of the lungs increased dramatically from an average of 1 for neo-435 mice to an average of 23 for SNCG-435 mice (Table 1). The representative lungs were shown in Fig. 2*C*. When these lungs were examined microscopically, large numbers of micrometastases were observed in SNCG-435 mice; the lungs of neo-435 mice had significantly fewer micrometastases (data not shown). Representative H&E-stained sections for neo-435 and SNCG-435 lungs are shown in Fig. 2, *D–G*. To our knowledge, human breast cancer cells usually do not form such a profound regional and metastatic tumor dissemination (visible lung nodules) in the spontaneous mammary fat pad nude mouse model. This dramatic SNCG-stimulated metastasis suggests a role for SNCG as a key positive regulator for breast cancer invasion and metastasis. The overexpression of SNCG in malignant infiltrating breast epithelial cells compared to the low expression level in noninvasive *in situ* carcinoma (2) suggests that SNCG expression is a meaningful marker for breast cancer malignant progression and may signal the more

Table 1 *Effects of SNCG on tumor incidence, tumor size, and axillary lymph node and lung metastasis*

Cells (400,000) were injected at day 1 into the mammary fat pads, and tumor volumes and lymph node and lung micrometastases were determined. Lymph node metastases were measured by microscopic examination for morphological evidence of tumor cells on the fixed axillary lymph nodes. Lung metastases were measured by the presence of visible tumor nodules on the surface of the lung. Volumes are expressed as the means ± SEs (number of tumors assayed). Experiment 1 had a total of 16 injections for eight mice in each group, and the mice were killed 42 days after injection. For experiment 2, there was a total of 10 injections for five mice in each group, and the mice were killed 38 days after injection. Statistical comparisons for SNCG-positive clones and SNCG-negative clones showed that there was no significant difference in the mean tumor sizes between pooled SNCG-positive and pooled SNCG-negative tumors. The lymph node positivity of pooled SNCG-435-1 tumors *versus* combined pooled SNCG-negative neo-435-1 and neo-435-2 tumors was $P < 0.039$ and $P < 0.029$ for pooled SNCG-435-3 tumors *versus* SNCG-negative tumors. Statistical comparison of primary tumors was analyzed by Student's t test. A χ^2 test was used for a statistical analysis of lymph node metastasis.

Experiment	Clones	Volume (cm^3) of the primary tumor	Tumor incidence	Lymph node	Lymph node	Lung metastasis
			Tumor total (%)	Average weight (mg)	No. positive/total no.	No. of nodules
1	neo-435-1	1.74 ± 0.52	16/16 (100)	14	3/16 (19)	0
	neo-435-2	1.9 ± 0.31	14/16 (88)	18	4/15 (27)	2
	SNCG-435-1	1.45 ± 0.37	15/16 (94)	43	10/15 (67)	19
	SNCG-435-3	1.78 ± 0.31	16/16 (100)	50	12/16 (75)	31
2	neo-435-1	1.35 ± 0.39	9/10 (90)	12	3/10 (30)	1
	neo-435-2	1.69 ± 0.44	10/10 (100)	15	3/9 (33)	1
	SNCG-435-1	1.73 ± 0.45	10/10 (100)	45	6/10 (60)	24
	SNCG-435-3	1.49 ± 0.34	10/10 (100)	39	7/9 (78)	17

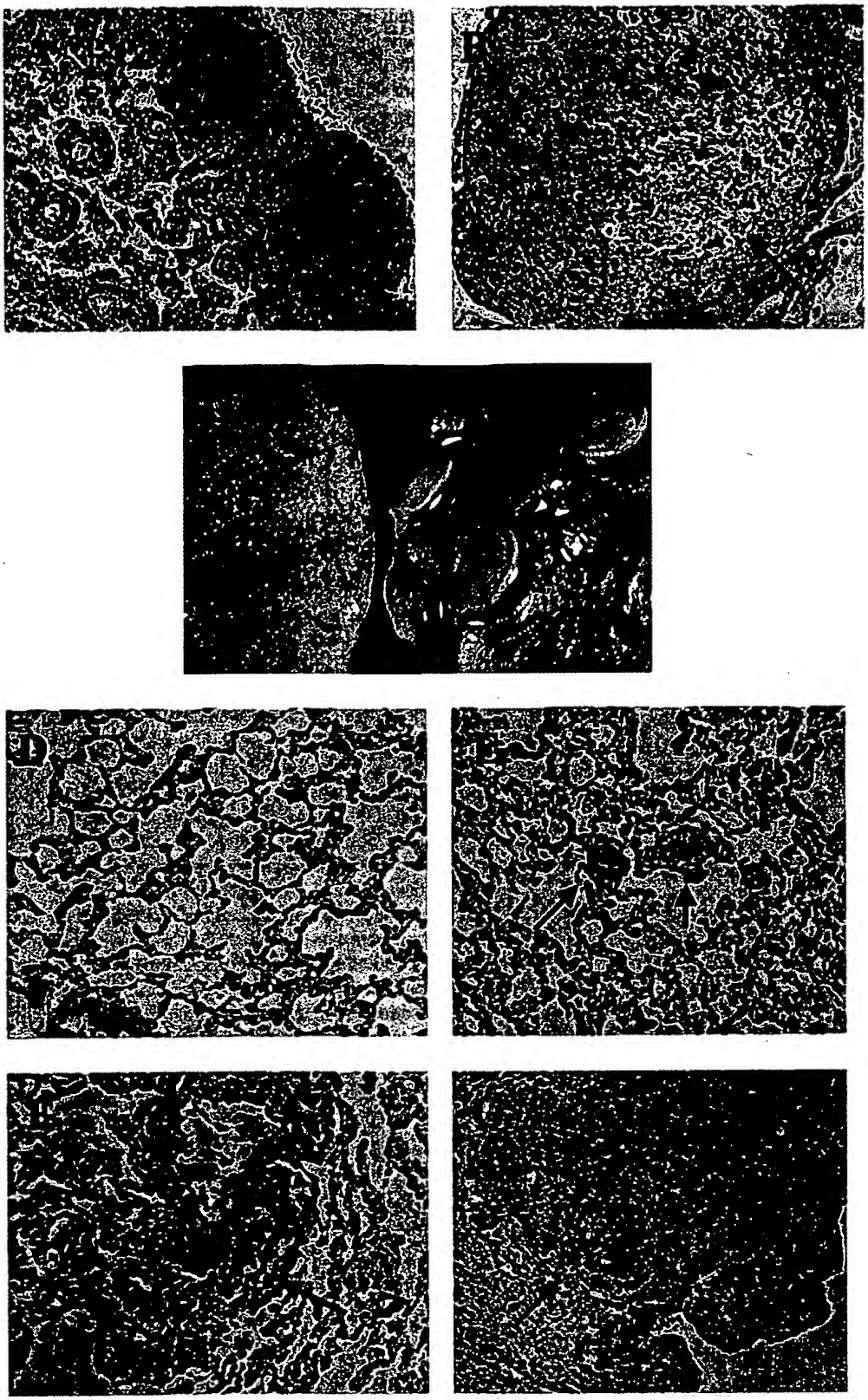


Fig. 2. Axillary lymph nodes and lung metastasis from neo-435 mice and SNCG-435 mice. The mice were sacrificed at day 40 after cell injection. Lymph nodes and lungs were isolated, and some were subjected to H&E staining. Representative axillary lymph nodes from a neo-435-1 mouse (*A*) and a SNCG-435–3 mouse (*B*) are shown. *Arrow*, an invasive breast tumor that mainly occupied the lymph node in a SNCG-435-3 mouse. *A* and *B*, ×10. *C*, representative lung metastases from mice injected with SNCG-positive and SNCG-negative cells. The *left lung* was from a neo-435-1 mouse, and the *right lung* was from a SNCG-435-3 mouse. The metastatic tumors only reflect the nodules on the surface of the lungs (×2.5). *D–G*, microscopic examination of representative lung metastases in H&E-stained sections. *D*, a lung without metastases from a neo-435-1 mouse. *E*, a lung with micrometastases from a neo-435-2 mouse. *F*, a lung with a small breast tumor nodule from a SNCG-435-1 mouse. *G*, a lung with a large breast tumor nodule from a SNCG-435-3 mouse. *Arrows*, breast tumors or cancer cells. *D–G*, ×20.

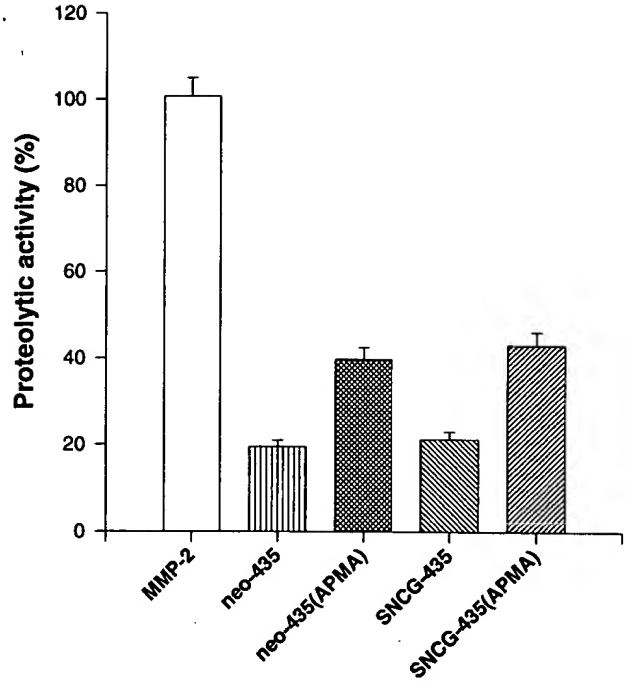


Fig. 3. Analysis of the MMP activities of SNCG-positive and SNCG-negative cells. The pooled CM from SNCG-negative neo-435-1 and neo-435-2 cells and SNCG-positive SNCG-435-1 and SNCG-435-3 cells were collected, concentrated 5-fold, normalized for protein concentrations, and subjected to MMP activity analysis. Recombinant AMPA-activated MMP-2 (80 ng) was used as a positive control. All values were normalized to the enzymatic activity of the recombinant MMP-2, which was taken as 100%. The numbers represent the means ± SD of three tests.

advanced invasive/metastatic phenotype of human breast cancer. In this regard, the up-regulation of SNCG expression may facilitate breast cancer progression leading to metastasis.

MMP Activity. In an effort to investigate the molecular mechanisms underlying SNCG-induced metastasis, we studied several invasion-related factors, including MMP and cell motility. The amyloid protein has recently been demonstrated to be a strong stimulator of MMP-2 and MMP-9 expression in astrocytes (14). It is well established that the overproduction and unrestrained activity of MMPs, particularly MMP-2 and MMP-9, are linked to the malignant conversion of a variety of different tumor cells (15–22) including mammary tumors (18–22). It is interesting to test whether SNCG, an amyloid-related protein, stimulates MMP-2 and MMP-9 expression in breast cancer cells and leads to the more metastatic phenotype. We investigated whether SNCG overexpression would increase MMP activity in MDA-MB-435 cells. In this regard, the pooled CM from two SNCG-negative cells and the pooled CM from two SNCG-positive cells were concentrated and subjected to a MMP enzymatic assay. As shown in Fig. 3, no significant differences in the basal levels of proteolytic activities were observed between neo-435 and SNCG-435 clones. Mammalian MMPs are usually secreted as latent proenzymes (zymogen) and require activation for their enzymatic activity. The incubation of CM with the MMP activator organomercurial compound APMA resulted in an approximately 2-fold increase in proteolytic activity for the CM from both neo-435 and SNCG-435 clones. However, no significant difference in APMA-activated MMP activities was observed between neo-435 and SNCG-435 clones. Because the measured enzymatic activity represents the net MMP activity, reflecting the balance between activated MMPs and the tissue inhibitors of metalloproteinase, our data suggest that SNCG-induced metastasis may not be mediated by the regulation of MMP and tissue inhibitors of metalloproteinase.

Stimulation of Invasiveness and Motility of MDA-MB-435 Cells by SNCG. We used an *in vitro* reconstituted basement membrane (Matrigel) invasion assay to determine the effect of SNCG on cell invasion. All three SNCG-negative cells (parental MDA-MB-435, neo-435-1, and neo-435-2) were moderately invasive. At the end of a 48-h incubation, an average of approximately 250 SNCG-negative cells had crossed the Matrigel barrier. A significant stimulation of invasiveness was noted in two SNCG-positive clones, with a 3-fold increase for SNCG-435-1 cells and a 4.3-fold increase for SNCG-435-3 cells (Fig. 4*A*). We also investigated the effect of SNCG on cell migration without Matrigel. A similar SNCG-stimulated pattern of migration was observed. At the end of a 24-h incubation, SNCG-435-1 cells migrated 4-fold, and SNCG-435-3 cells migrated 4.2-fold over that of average SNCG-negative cells (Fig. 4*B*). The similar magnitude of the invasion-stimulating and migration-stimulating activity of SNCG suggests that the increased invasion in SNCG clones may be mediated by an alteration of cell motility. To determine whether the increased cell motility is mediated by chemotaxis due to the different concentrations of serum or chemoattractants in the top and bottom chambers, we compared the migration of SNCG-435-3 and neo-435-1 cells under three different culture conditions: (*a*) serum-free conditions; (*b*) serum with gradient; and (*c*) serum without gradient. As shown in Fig. 5, although the migration was relatively

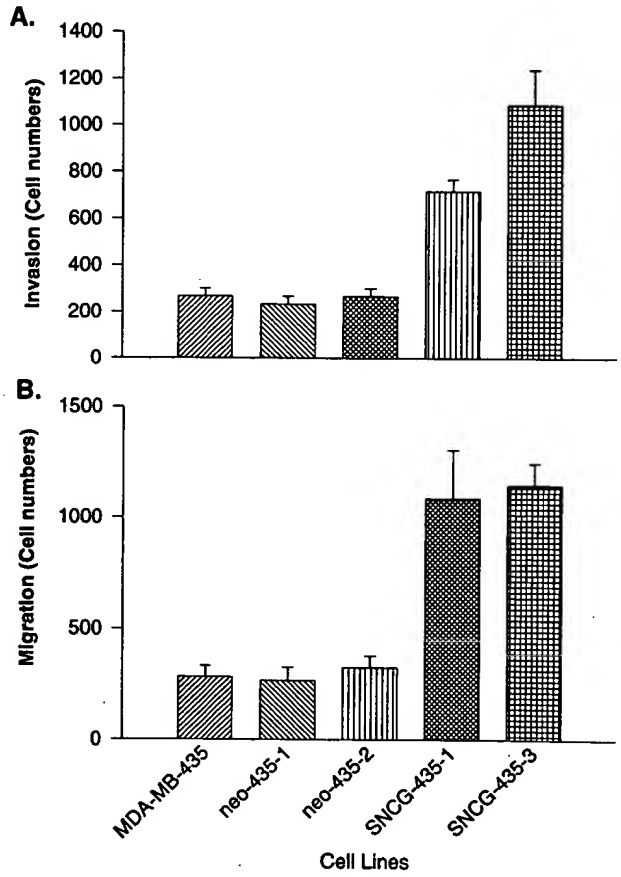


Fig. 4. Stimulation of invasiveness and migration of MDA-MB-435 cells by SNCG. Cells were seeded at a density of 30,000 cells/ml/well on 8-μm polycarbonate membranes coated with (*A*) or without (*B*) 4 mg/ml growth factor-reduced Matrigel. The top chamber contained 5% FCS, and the bottom chamber contained 10% FCS. *A*, after incubation in a humidified incubator with 5% CO_2 at 37°C for 48 h, the medium and cells were removed from the bottom chambers and counted using a microscope. *B*, cells were cultured under the same conditions as described in *A*. The number of cells that migrated was counted after a 24-h incubation. All values were expressed as the number of invaded cells. The numbers represent the means ± SD of three cultures.

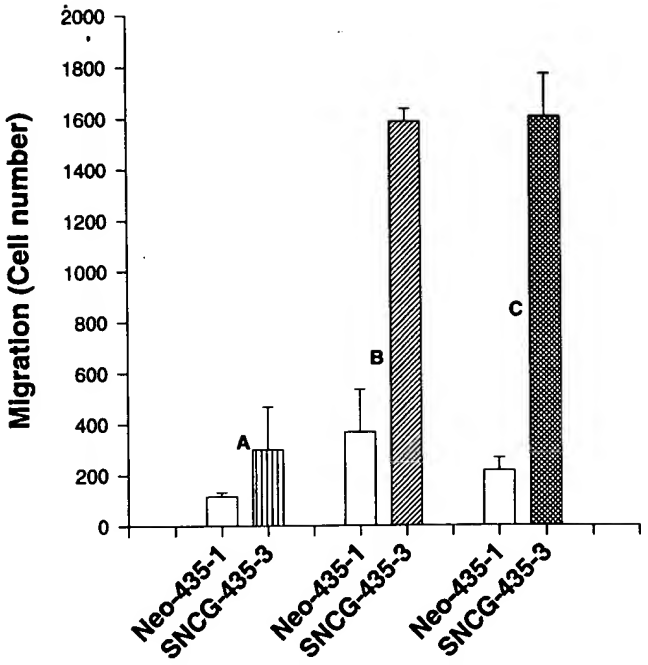


Fig. 5. Comparison of the cell migration of SNCG-435-3 and neo-435-1 cells under different conditions. Cells were cultured on noncoated membrane at a density of 30,000 cells/ml/well. The cells that migrated were harvested at 32 h after incubation. *A*, 0% serum in both the top and bottom chambers. *B*, 2% serum in both the top and bottom chambers. *C*, 2% serum in the top chamber and 10% serum in the bottom chamber. All values were expressed as the number of invaded cells. The numbers represent the means ± SD of triplet wells.

low under serum-free conditions, there was a 2.8-fold increase in migration in SNCG-435-3 cells compared with that in neo-435-1 cells. When 2% serum was added in the top chamber, the migration of both SNCG-positive and -negative cells increased significantly. However, the migration of SNCG-435-3 cells was not affected by the serum gradient. Approximately 1600 of the SNCG-435-3 cells that migrated into the bottom chamber contained either 2% serum or 10% serum. These data suggest that the increased migration in SNCG-positive cells is not likely to be mediated by chemotaxis but rather by high motility features intrinsic to the cells.

Many breast tumors go through a series of events from the time of initial detection to the formation of the lethal invasive and metastatic stage. According to the three-step hypothesis of invasion (23), cell adhesion, local proteolysis, and subsequent migration or motility are key steps in the traversal of the basement membrane and connective tissue. In this study, we provide evidence linking the overexpression of neural protein SNCG, a previously identified breast cancer-specific gene (2), in human breast cancer cells with increased motility and invasive activity *in vitro* and a profound augmentation of metastasis *in vivo*.

SNC proteins have a structural resemblance to apolipoproteins but are abundant in the neuronal cytosol and are present in enriched amounts at presynaptic terminals (9). SNCs have been specifically implicated in two diseases: AD and PD. In AD patients, a peptide derived from SNCA forms an intrinsic component of plaque amyloid (9). In PD patients, a SNCA allele is genetically linked to several independent familial cases, and the protein appears to accumulate in Lewy bodies (9). The general significance of the involvement of neural protein SNCG in cancer metastasis is unknown. Recently, SNCA and SNCB were identified as two abundant proteins through their reactivity with a monoclonal antibody recognizing MAP-τ (6) on immunoblots. In eukaryotic cells, microtubules, actin, and intermediate filaments interact to form the cytoskeletal network involved in the determination of cell architecture, mitosis, differentiation, and motility (24). Cytoskeletal organization and dynamics depend on protein self-associations and interactions with regulatory elements such as MAPs. There is increasing evidence that MAPs, including MAP-τ, play a critical role in inducing microtubule assembly and controlling the dynamic instability of microtubules, thus controlling the state of their assembly and organization in cells (reviewed in Ref. 24). SNCG may interact with MAPs and regulate the cytoskeletal organization and dynamics, leading to increased motility. Nevertheless, our data indicate that the increased expression of SNCG correlates with breast cancer progression (2) and leads to a more malignant metastatic phenotype. We also demonstrated that SNCG expression in breast cancer cells was subjected to cytokine regulation and dramatically suppressed by the tumor growth inhibitor OM, and that this OM-induced transcriptional suppression of the SNCG gene was associated with OM-mediated growth inhibition.[4] OM is an antitumor cytokine produced mainly by activated T cells and macrophages (25), and its growth-suppressing activity has been well studied in breast cancer cells (26–28). One characteristic of the host response to tumor progression is the infiltration of tumors by macrophages and T lymphocytes. The production of tumor-suppressing cytokines in a timely and locally (*in situ*) released fashion may represent an important function of the host defense system in suppressing tumor progression. From this prospective view, the dramatic suppression of SNCG expression in malignant breast cells by OM may represent the host-mediated tumor suppression leading to the inhibition of breast cancer progression.

This is the first report indicating the potential involvement of SNC in a non-neural disease. An elucidation of the reasons for SNCG overexpression in infiltrating breast cancer and SNCG-induced metastasis may shed some light on the pathogenesis of breast cancer progression as well as neurodegenerative disorders.

REFERENCES

1. Porter-Jordan, K., and Lippman, M. E. Overview of the biological markers of breast cancer. Hematol. Oncol. Clin. N. Am., *8:* 73–100, 1994.
2. Ji, H., Liu, Y. E., Jia, T., Wang, M., Liu, J., Xiao, G., Joseph, B. K., Rosen, C., and Shi, Y. E. Identification of a breast cancer-specific gene, *BCSG1*, by direct differential complementary DNA sequencing. Cancer Res., *57:* 759–764, 1997.
3. Shi, Y. E., Ni, J., Xiao, G., Liu, Y. E., Fuchs, A., Yu, G., Su, J., Cosgrove, J. M., Xing, L., Zhang, M., Li, J., Aggarwal, B. B., Meager, A., and Gentz, R. Antitumor activity of the novel human breast cancer growth inhibitor, mammary-derived growth inhibitor-related gene, *MRG*. Cancer Res., *57:* 3084–3091, 1997.
4. Ueda, K., Fukushima, H., Masliah, E., Xia, Y., Iwai, A., Yoshimoto, M., Otero, D. A., Kondo, J., Ihara, Y., and Saitoh, T. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proc. Natl. Acad. Sci. USA, *90:* 11282–11286, 1993.
5. Xia, Y., Rohan-de-Silva, H. A., Rosi, B. L., Yamaoka, L. H., Rimmler, J. B., Pericak-Vance, M. A., Ross, A. D., Chen, X., Masliah, E., DeTeresa, R., Iwai, A., Sundsmo, M., Thomas, R. G., Hofstetter, C. R., Gregory, E., Hansen, L. A., Katzman, R., Thal, L. J., and Saitoh, T. Genetic studies in Alzheimer's disease with an NACP/α-synuclein polymorphism. Ann. Neurol., *40:* 207–215, 1996.
6. Jakes, R., Spillantini, M. G., and Goedert, M. Identification of two distinct synucleins from human brain. FEBS Lett., *345:* 27–34, 1994.
7. Spillantini, M. G., Schmidt, M. L., Lee, V. M., Trojanowski, J. Q., Jakes, R., and Goedert, M. α-Synuclein in Lewy bodies. Nature (Lond.), *388:* 839–840, 1997.
8. Masters, C. L., Simms, G., Wenman, N. A., Multhup, G., McDonald, B. L., and Beyreuther, K. Amyloid plaque core in Alzheimer's disease and Down syndrome. Proc. Natl. Acad. Sci. USA, *82:* 4245–4249, 1985.
9. Clayton, D. F., and George, J. M. The synucleins: a family of proteins involved in synaptic function, plasticity, neurodegeneration and disease. Trends Neurosci., *21:* 249–254, 1998.
10. Takeda, A., Mallory, M., Sundsmo, M., Honer W., Hansen, L., and Masliah, E. Abnormal accumulation of NACP/α-synuclein in neurodegenerative disorders. J. Pathol., *152:* 367–372, 1998.

[4] J. Liu, M. J. Spence, Y. L. Zhang, T. Jia, Y. E. Liu, and Y. E. Shi. Transcriptional suppression of breast cancer-specific SNCG expression by the growth-inhibitory cytokine OM, submitted for publication.

11. Wang, M., Liu, Y. E., John, G., Sheng, S., Fuchs, F., Rosen, E. M., and Shi, Y. E. Inhibition of tumor growth and metastasis of human breast cancer cells transfected with tissue inhibitor of metalloproteinase 4. Oncogene, *14:* 2767–2774, 1997.
12. Liu, Y. E., Wang, M., Greene, J., Su, J., Ullrich, S., Sheng, S., Alexander, P., Sang, Q. A., and Shi, Y. E. Preparation and characterization of recombinant TIMP-4 protein. J. Biol. Chem., *272:* 20479–20483, 1997.
13. Shi, Y. E., Torri, J., Yieh, L., Wellstein, A., Lippman, M. E., and Dickson, R. B. Isolation and characterization of a novel matrix-degrading protease from hormone-dependent human breast cancer cells. Cancer Res., *53:* 1409–1415, 1993.
14. Deb, S., and Gottschall, P. E. Increased production of matrix metalloproteinases in enriched astrocyte and mixed hippocampal cultures treated with β-amyloid peptides. J. Neurochem., *66:* 1641–1647, 1996.
15. Pyke, C., Ralfkiaer, E., Tryggvason, K., and Dano, K. Messenger RNA for two type IV collagenases is located in stromal cells in human colon cancer. Am. J. Pathol., *142:* 359–365, 1993.
16. Naylor, M. S., Stamp, G. W., Davies, B. D., and Balkwill, F. R. Expression and activity of MMPS and their regulators in ovarian cancer. Int. J. Cancer, *58:* 50–56, 1994.
17. Polette, M., Clavel, C., Birembaut, P., and De-Clerck, Y. A. Localization by *in situ* hybridization of mRNAs encoding stromelysin 3 and tissue inhibitors of metalloproteinases TIMP-1 and TIMP-2 in human head and neck carcinomas. Pathol. Res. Pract., *189:* 1052–1057, 1993.
18. Urbanski, S. J., Edwards, D. R., Maitland, A., Leco, K. J., Watson, A., and Kossakowska, A. E. Expression of metalloproteinases and their inhibitors in primary pulmonary carcinomas. Br. J. Cancer, *66:* 1188–1194, 1992.
19. Tryggvason, K., Hoyhtya, M., and Pyke, C. Type IV collagenases in invasive tumors. Breast Cancer Res. Treat., *24:* 209–218, 1993.
20. Polette, M., Clavel, C., Cockett, M., Girod-de-Bentzmann, S., Murphy, G., and Birembaut, P. Detection and localization of mRNAs encoding matrix metalloproteinases and their tissue inhibitor in human breast pathology. Invasion Metastasis, *13:* 31–37, 1993.
21. Basset, P., Wolf, C., Rouyer, N., Bellocq, J. P., Rio, M. C., and Chambon, P. Stromelysin-3 in stromal tissue as a control factor in breast cancer behavior. Cancer (Phila.), *74:* 1045–1049, 1994.
22. Heppner, K. J., Matrisian, L. M., Jensen, R. A., and Rodgers, W. H. Expression of most matrix metalloproteinase family members in breast cancer represents a tumor-induced host response. Am. J. Pathol., *149:* 273–282, 1996.
23. Stettler-Stevenson, W. G., Aznavoorian, S., and Liotta, L. A. Tumor cell interactions with the extracellular matrix during invasion and metastasis. Annu. Rev. Cell Biol., *9:* 541–573, 1993.
24. Maccioni, R. B., and Cambiazo, V. Role of microtubule-associated proteins in the control of microtubule assembly. Physiol. Rev., *75:* 835–865, 1995.
25. Zading, J. M., Shoyab, M., Marquardt, H., Hanson, M., Lioubin, M. N., and Tdaro, G. J. Oncostatin M: a growth regulator produced by differentiated lymphoma cells. Proc. Natl. Acad. Sci. USA, *83:* 9739–9743, 1986.
26. Liu, J., Spence, M. J., Wallace, P. M., Forcier, K., Hellstrom, I., and Vestal, R. E. Oncostatin M-specific receptor mediates inhibition of breast cancer cell growth and down-regulation of the c-*myc* proto-oncogene. Cell Growth Differ., *8:* 667–676, 1997.
27. Spence, M. J., Vestal, R. E., and Liu, J. Oncostatin M-mediated transcriptional suppression of the c-*myc* gene in breast cancer cells. Cancer Res., *57:* 2223–2228, 1997.
28. Douglas, A. M., Grant, S. L., Goss, G. A., Clouston, D. R., Sutherland, R. L., and Begley, C. G. Oncostatin M induces the differentiation of breast cancer cells. Int. J. Cancer, *75:* 64–73, 1998.

[CANCER RESEARCH 57, 759–764, February 15, 1997]

Identification of a Breast Cancer-specific Gene, *BCSG1*, by Direct Differential cDNA Sequencing[1]

Hongjun Ji, Yiliang E. Liu, Tongli Jia, Mingsheng Wang, Jingwen Liu, Guowei Xiao, Benjamin K. Joseph, Craig Rosen, and Y. Eric Shi[2]

Departments of Pediatrics [Y. E. L., T. J., M. W., G. X., B. K. J., Y. E. S.] and Pathology [Y. E. S.], Long Island Jewish Medical Center, The Long Island Campus for the Albert Einstein College of Medicine, New Hyde Park, New York 11040; Human Genome Sciences, Inc., Rockville, Maryland 20850-3338 [H. J., C. R.]; and Mountain States Medical Research Institute and Department of Veterans Affairs Medical Center, Boise, Idaho 83702 [J. L.]

ABSTRACT

A high-throughput direct-differential cDNA sequencing approach was employed to identify genes differentially expressed in normal breast as compared with breast cancer. Approximately 6000 expressed sequence tags (ESTs) from cDNA libraries of normal breast and breast carcinoma were selected randomly and subjected to EST-sequencing analysis. The relative expression levels of more than 2000 unique EST groups were quantitatively compared in normal *versus* cancerous breast. Of many putative differentially expressed genes, a breast cancer-specific gene, *BCSG1*, which was expressed in high abundance in a breast cancer cDNA library but scarcely in a normal breast cDNA library, was identified as a putative breast cancer marker. *In situ* hybridization analysis demonstrated stage-specific BCSG1 expression as follows: BCSG1 was undetectable in normal or benign breast lesions, showed partial expression in ductal carcinoma *in situ*, but was expressed at an extremely high level in advanced infiltrating breast cancer. The predicted amino acid sequence of *BCSG1* gene has a significant sequence homology to the non-amyloid β protein fragment of the Alzheimer's disease amyloid protein. BCSG1 overexpression may indicate breast cancer malignant progression from benign breast or *in situ* carcinoma to the highly infiltrating carcinoma.

INTRODUCTION

The onset and progression of breast cancer is accompanied by multiple genetic changes that result in qualitative and quantitative alterations in individual gene expression (1). Our hypothesis is that many of these quantitative genetic changes manifest themselves as alterations in the cellular complement of novel transcribed mRNAs. Identification of these mRNAs, if sufficiently characterized, could provide clinically useful information for patient management and prognosis while enhancing our understanding of breast cancer pathogenesis. Although pathological end points such as tumor size, lymph node status, and status of estrogen receptor and progesterone receptor remain the most useful guides in prognosis and in selecting treatment strategies for breast cancer (2), there is a need to further investigate the molecular mechanisms that determine the properties of an individual tumor, *e.g.*, probability of metastasis. Although numerous prognostic factors have now been identified, few have contributed to defining the clinical response to therapy.

Identification of quantitative changes in gene expression that occur in the malignant mammary gland, if sufficiently characterized, may yield novel molecular markers that may be useful in the diagnosis and treatment of human breast cancer. Several differential cloning methods, such as differential display PCR and subtractive hybridization, have been used to identify the genes differentially expressed in breast cancer biopsies, as compared to normal breast tissue controls (3–7). However, these investigations have involved the relatively time- and labor-intensive steps of subcloning, library screening, and cDNA sequencing of individual genes (4, 8). On the other hand, creation of libraries is a rapid method used to identify or "tag" sequences that are expressed in specific tissues (9, 10). Since the introduction of the EST[3] sequencing approach, many novel human genes have been discovered (9, 10). The advantage of this methodology, compared to isolation and sequencing of individual cDNAs, is that a large number of sequences can be "catalogued" with small amounts of sequencing data.

With the availability of tens of thousands of ESTs, researchers now shift their attention to the unveiling of the expression profile of individual genes or patterns of genes in normal *versus* diseased states. Several newly developed strategies, such as the serial analysis of gene expression (11) and cDNA microarray (12) methods, have demonstrated potential for broad application for quantitative analysis of differential patterns of gene expression. Within this context, we undertook a search, using the differential cDNA sequencing approach, for isolation of differentially expressed ESTs and the possible presence of the new marker genes for breast cancer. In this initial report, we describe a novel BCSG named *BCSG1* that is overexpressed in advanced infiltrating breast cancer cells but not in normal or benign breast lesion. The expression pattern of BCSG1 may be a meaningful marker in the development of breast cancer.

MATERIALS AND METHODS

Reagents. Restriction enzymes, T7 polymerase, random primer DNA labeling kit, and digoxigenin-labeled nucleotides were obtained from Boehringer Mannheim (Indianapolis, IN). [^{32}P]dATP was purchased from Amersham Corp.

Differential cDNA Sequencing. We have used EST analysis to search for new genes differentially expressed in breast cancer *versus* normal breast tissue. A data base containing approximately 500,000 human partial cDNA sequences (ESTs) has been established in a collaborative effort between the Institute for Genomic Research and Human Genome Sciences, Inc., using high-throughput automated DNA sequence analysis of randomly selected human cDNA clones (10). RNAs from a stage III breast carcinoma and patient-matched normal breast were isolated and subjected to preparation of cDNA libraries. EST-automated DNA sequence analysis was performed on randomly selected cDNA clones. Both libraries had about 60% novel gene sequences, which did not match exactly to published human genes. A total of 3048 ESTs from breast cancer cDNA library and 2886 ESTs from the normal breast cDNA library were randomly picked and sequence analyzed. The ESTs with overlapping sequences were grouped into unique EST groups, with each EST group representing a gene or a family of sequence-related genes. Each unique EST group without overlapping sequences was analyzed for its relative expression by examining the number of expressed individual ESTs in the libraries of normal *versus* diseased tissues. There were more than 2200 EST groups that were analyzed for quantitative comparison of EST "hits" in the pair of cDNA

Received 8/22/96; accepted 12/20/96.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

[1] Supported in part by grants from US Army Breast Cancer Research Program (DAMD17-94-J-4149) and Helen and Irving Schneider.

[2] To whom requests for reprints should be addressed, at Pediatric Research Center, Schneider Children's Hospital, Long Island Jewish Medical Center, Albert Einstein College of Medicine, New Hyde Park, NY 11042. Phone: (718) 470-3086; Fax: (718) 470-6744; E-mail: shi@aecom.yu.edu.

[3] The abbreviations used are: EST, expressed sequence tag; BCSG, breast cancer-specific gene; Aβ, amyloid β protein; AD, Alzheimer's disease; DCIS, ductal carcinoma *in situ*.

Table 1 *Partial list of differentially expressed genes in normal versus cancerous breasts identified by differential cDNA sequencing*

Complementary DNA libraries were established from a stage III breast carcinoma and patient-matched normal breast. A total of 5934 ESTs were randomly picked and sequence analyzed. More than 2200 distinctive EST groups were analyzed for quantitative comparison of EST hits in the pair of cDNA libraries from breast cancer *versus* normal breast as described in "Materials and Methods." The same EST groups were also analyzed by examining the tissue-specific expression against the total of 500,000 ESTs from a variety of different cDNA libraries. Only a unique EST group with more than three breast-specific EST hits was listed, and the rest of the several dozen EST groups with fewer than four breast-specific EST hits were omitted in this list.

Genes more abundant in breast cancer

	ESTs	
Class I Genes	Cancer	Normal
Breast basic conserved gene	33	9
Cathepsin D	5	1
M_r 67,000 laminin receptor	4	0
Elongation factor 1	13	5

Genes more abundant in normal breast

	ESTs	
Class II Genes	Cancer	Normal
Matrix Gla protein	0	8
M_r 23,000 highly basic protein	3	11

Genes as breast-specific and differentially expressed

	ESTs		
Class III Genes	NB[a]	BC[b]	All tissues
BCSG1	1	6	8[c]
BCSG2	0	7	7
BCSG3	0	5	5
BCSG4	4	0	4
BCSG5	0	4	4

[a] normal breast; [b] breast cancer; [c] seven ESTs from breast libraries and one EST from brain library.

libraries from normal breast *versus* breast cancer by examining the expression of individual EST sequences. The number of EST hits in the libraries reflects the relative expression or mRNA transcript copy numbers of the EST. This direct differential cDNA sequence, utilizing the direct EST sequencing analysis simultaneously on a pair of cDNA libraries made from normal breast and breast cancer tissue, was used to study the expression profile of individual genes and patterns of genes in normal breast *versus* breast cancer tissue.

Tissue-specific Expression Analysis. Analysis of relative expression of breast-derived ESTs *versus* their expression in other tissues was performed. The differentially expressed EST groups identified by differential cDNA sequence were analyzed for tissue-specific expression against the total of 500,000 ESTs from a variety of different cDNA libraries.

Northern Analysis. Total RNA was extracted from tissues according to the method of Chomczynski and Sacchi (13). The RNA from human breast cancer cells was prepared using the RNA isolation kit RNAzol B (Tel-Test, Inc.) based on the manufacturer's instruction. Equal aliquots of RNA were electrophoresed in a 1.2% agarose gel containing formaldehyde and transferred to nylon membrane (Boehringer Mannheim). The membrane was prehybridized with ExpressHyb hybridization solution (Clontech, Inc.) at 68°C for 30 min. The hybridization was carried out in the same solution with ^{32}P-labeled BCSG1 probe (1.5×10^6 cpm/ml) for 1 h at 68°C. The membrane was then rinsed in 2× SSC containing 0.05% SDS three times for 30 min at room temperature, followed by two washes with 0.1× SSC containing 0.1% SDS for 40 min at 50°C. The full-length BCSG1 cDNA was isolated from the Bluescript vector following *Eco*RI and *Xho*I digestion and was used as a template for preparation of a random-labeled cDNA probe.

***In Situ* Hybridization.** *In situ* hybridization was carried out as described (14). Briefly, deparaffinized and acid-treated sections (5 μm thick) were treated with proteinase K, prehybridized, and hybridized overnight with digoxigenin-labeled antisense transcripts from a BCSG1 cDNA insert. The BCSG1 antisense probe is a 550-bp full-length fragment. The probe was generated by a *Pst*I cut of BCSG1 cDNA plasmid and followed by T7 polymerase. Hybridization was followed by RNase treatment and three stringent washings. Sections were incubated with mouse antidigoxigenin antibodies (Boehringer Mannheim) followed by the incubation with biotin-conjugated secondary rabbit antimouse antibodies (DAKO). The colorimetric detections were performed with a standard indirect streptavidin-biotin immunoreaction method using the Universal LSAB Kit (DAKO) according to the manufacturer's instructions.

RESULTS

Molecular Cloning of BCSG1 cDNA. We generated cDNA libraries from breast cancer biopsy specimens and patient-matched normal breast and analyzed these libraries by EST sequencing. Approximately 6000 ESTs were analyzed and assigned to different groups based on sequence overlapping, and 2200 unique EST groups were first analyzed for relative expression in the cDNA libraries from normal breast *versus* breast cancer tissue and then subjected to tissue-specific expression by examining the tissue origins of individual EST sequences against a large population of ESTs derived from a variety of different tissue types. As a result, we identified three classes of EST groups that were differentially expressed in normal breast *versus* breast cancer tissue. As a demonstration of this approach, Table 1 shows a partial list of three classes of genes that are differentially expressed in normal breast *versus* breast cancer tissue. Class I represents the genes more abundant in breast cancer than in normal breast and includes cathepsin D, a well-studied steroid regulated extracellular matrix-degrading proteinase (15–17). Cathepsin D is thought to play a role in breast cancer metastasis (15–17) and has been proposed as a prognostic marker in breast cancer progression (18–21). As listed, there were five cathepsin D ESTs sequenced in the breast cancer cDNA library and only one EST in the normal breast cDNA library. Another proposed breast cancer metastasis-related gene and a prognostic marker for breast cancer, M_r 67,000 laminin receptor (22–26), was also picked up in this class by the differential cDNA sequencing approach. Class II represents genes that are more abundant in normal breast than in breast cancer.

```
1    M D V F K K G F S I A K K G V V G A V E    BCSG1
1    M D V F M K G L S K A K E G V V A A A E    *Human AD Amyloid

21   K T K Q G V T E A A E K T K E G V M Y V    BCSG1
21   K T K Q G V A E A A G K T K E G V L Y V    *Human AD Amyloid

41   G A K T Y E N V V Q S V T S V A E K T K    BCSG1
41   G S K T K E G V V H G V A T V A E K T K    *Human AD Amyloid

61   E Q A N A V S K A V V S S V N T V A T K    BCSG1
61   E Q V T N V G G A V V T G V T A V A Q K    *Human AD Amyloid

81   T V E E A E N I A V T S G V V R K E D L    BCSG1
81   T V E G A G S I A A A T G F V K K D Q L    *Human AD Amyloid

101  R P S A P Q Q E G E A S K E K E E V A E    BCSG1
101  G K N E E G A P Q E G I L E D M P V D P    *Human AD Amyloid

121  E A Q S G G D                             BCSG1
121  D N E A Y E M P S E E G Y Q D Y E P E A    *Human AD Amyloid
```

*Non-Aß component of Alzheimer's disease (AD) Amyloid

Fig. 1. Comparison of the predicted amino acid sequence with the sequence of non-Aβ component of AD amyloid protein using SwissProt. After optimal alignment using the clustal method of the MegAlign Program from the DNASTAR software package, the putative protein shows a 54% sequence identity with the non-Aβ fragment of human AD amyloid protein. Conserved amino acids are underlined.

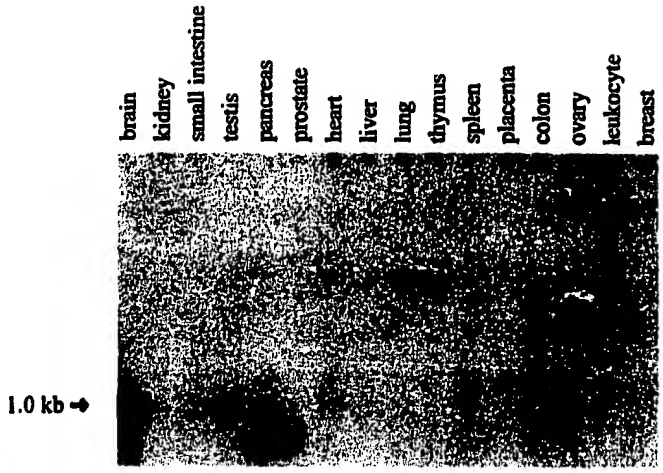


Fig. 2. The expression of *BCSG1* gene in a variety of normal human adult tissues. Twenty μg of total RNA from each of the above tissues were analyzed in Northern blot using a random primer probe. A strong hybridizing band of about 1 kb was recognized in the lane corresponding to RNA from adult brain. A weak 1-kb transcript was also detected in testis, heart, spleen, colon, and ovary.

Although the genes in classes I and II are differentially expressed in normal breast *versus* breast cancer tissue, these genes are not unique to breast tissues. Class III is a special group of genes that are selectively expressed in breast relative to other tissue types. The tissue-specific expression of the unique gene was searched against approximately 500,000 ESTs using the BLAST program (27). None of these BCSGs except the first one matched with any sequences in public gene sequence databases. The automated screening revealed a group of eight ESTs encoding a novel *BCSG1* gene from the partial cDNA database containing approximately 500,000 ESTs. Of the eight distinctive EST clones in BCSG1, seven of them were discovered in breast cDNA libraries and only one in a brain library. Of the seven EST clones discovered in the breast cDNA libraries, six of them were identified in the breast tumor library and only one in the normal breast library. BCSG1 was chosen for analysis as a first putative breast cancer marker gene because (*a*) its sequence has been matched with the sequence in the public gene sequence database; and (*b*) most of the individual EST sequences in BCSG1 were derived from a breast tumor cDNA library. After sequencing analysis of all six EST clones derived from the breast cancer library, one EST clone was found to have a complete full-length sequence. The open reading frame of the resulting full-length gene is predicted to encode a 127-amino acid polypeptide. Comparison of the predicted amino acid sequence with the sequence of a similar human protein is shown in Fig. 1. After optimal alignment, the putative BCSG1-encoded protein shows 54% sequence identity with the recently cloned non-Aβ fragment of human AD amyloid protein (28).

Tissue Expression. The expression of *BCSG1* gene in a variety of normal human tissues were analyzed by Northern blotting (Fig. 2). As expected, the Northern blot showed that BCSG1 was abundantly expressed as a 1-kb transcript in brain, which is the rich source for the *AD* amyloid family gene. Similar bands with much lower accumulations in their relative intensities were also obtained in ovary, testis, colon, and heart. By contrast, none of them was present in other specimens analyzed, such as breast, kidney, liver, prostate, lung, small intestine, thymus, and placenta.

[4] J. Liu, M. J. Spence, P. M. Wallace, K. Forcier, I. Hellstrom, and R. T. Vestal. Oncostatin M-specific receptor mediates inhibition of breast cancer cell growth, antagonism of growth factors, and down regulation of *c-myc* proto-oncogene, submitted for publication.

Expression of BCSG1 in Human Breast Cancer Cells. In an attempt to evaluate the potential biological significance of BCSG1 on human breast cancer development and progression, we studied *BCSG1* gene expression in human breast cancer cells. Northern blot (Fig. 3) detected the 1-kb BCSG1 transcript in two of four lines derived from pleural effusion and four of four lines detected from ductal infiltrating carcinomas. Among these lines, H3922 expressed the highest level of BCSG1 mRNA. The absence of BCSG1 mRNA in some breast cancer cell lines may suggest that the expression of *BCSG1* gene requires specific *in vivo* conditions, or that it is induced by interactions between the tumor cells and stromal cells.

To localize the cellular source of the BCSG1 expression and to further assess the biological relevance of the overexpression of BCSG1 in breast cancers, we next performed *in situ* hybridization on fixed breast sections from 20 infiltrating carcinomas, 15 DCISs, and 18 benign breast lesions, including 5 reduction mammoplasty specimens, 8 breast hyperplasias, and 5 fibroadenomas. In these experiments, we examined two aspects of BCSG1 expression, including the tissue localization (stromal *versus* epithelial) and the correlation of BCSG1 expression and breast cancer malignant phenotype. There was a wide variation in staining intensity for BCSG1 expression among the human breast cancer specimens. Because the colorimetric *in situ* hybridization is not quantitative, the tissue samples were classified into either positive or negative staining for BCSG1 expression; no attempt was made to differentiate the levels of expression of BCSG1 among positive-staining specimens. The negative cases were confirmed with at least two independent experiments. All stainings were reviewed by at least two people. Fig. 4 shows a representative *in situ* hybridization for BCSG1. We found a strongly positive BCSG1 hybridization in neoplastic epithelial cells of highly infiltrating breast carcinomas (Fig. 4, *A* and *B*). The expression of BCSG1 mRNA was detectable in the neoplastic epithelial cells in 17 of 20 infiltrating breast carcinomas. No expression of BCSG1 was detected in the stromal cells. In contrast, expression of BCSG1 was absent in 16 out of 18 cases of normal or benign breast lesions. A representative negative staining of BCSG1 in normal ductal breast epithelial cells (Fig. 4*E*), a benign proliferative breast lesion (Fig. 4*F*), and a benign fibroadenoma (Fig. 4*G*) are presented. Furthermore, as demonstrated in Fig. 4*B* for a highly invasive breast carcinoma, no detectable signal of BCSG1 expression was evident in the residual normal lobular breast epithelial cells, although the surrounding invasive breast carcinoma cells were stained positive for BCSG1 expression. The presence of BCSG1 transcript in human breast tissue and its overexpression in breast carcinomas are consistent with our differential cDNA sequencing cloning strategy, which suggests a possible

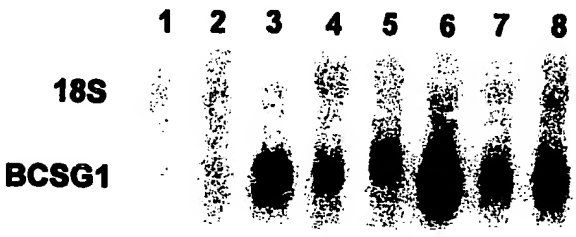


Fig. 3. Northern blot analysis of BCSG1 expression in human breast cancer cell lines. Total RNA was isolated and analyzed (20 μg/*Lane*) by Northern blot. After hybridization and washing, the filter was exposed to X-ray film for 48 h. The integrity and the loading control of the RNAs were ascertained by direct visualization of the 18 S rRNA in stained gel. *Lane 1*, H3396 (derived from pleural effusion); *Lane 2*, MCF7 (derived from pleural effusion); *Lane 3*, SKBR-3 (derived from pleural effusion); *Lane 4*, MDA-MB-231 (derived from pleural effusion); *Lane 5*, H3914 (derived from infiltrating ductal carcinoma); *Lane 6*, H3922 (derived from infiltrating ductal carcinoma); *Lane 7*, ZR-75–1 (derived from infiltrating ductal carcinoma); *Lane 8*, T47D (derived from infiltrating ductal carcinoma). Cell lines T47D, ZR-75–1, SKBR-3, MCF-7, and MDA-MB-231 are from American Type Culture Collection; all other lines were isolated initially at Bristol-Myers Squibb Pharmaceutical Research Institute.[4]

from virtually no detectable expression in normal or benign breast to partial expression (7 of 15) in the *in situ* breast carcinoma and to the high expression (17 of 20) in the infiltrating malignant breast carcinomas, suggest an association of BCSG1 expression with breast cancer malignant progression. On the basis of this BCSG1 expression pattern, we propose that BCSG1 may be used potentially as a breast cancer progression marker.

DISCUSSION

More than 190,000 new cases of breast cancer are diagnosed in the United States every year, with incidence increasing by approximately 1% annually (29, 30). Studies linked to the discovery of new genetic markers will provide new information leading to the understanding of breast cancer development and progression. There are two classes of genes affecting tumor development. Genes influencing the cancer phenotype that act directly as a result of changes (*e.g.*, mutation) at the DNA level, such as *BRCA1*, *BRCA2*, and *p53*, are called Class I genes. The Class II genes affect the phenotype by modulation at the expression level. Development of breast cancer and subsequent malignant progression is associated with alterations of a variety of genes of both classes. Many new predictive and prognostic factors have been proposed and studied for breast cancer. HER 2/neu-positive tumors respond poorly to endocrine treatment (31, 32). p53 alteration has an indication of poorer prognosis and poor response to tamoxifen (33, 34). The lack of Nm23 expression has an indicative value of metastatic potential and poor prognosis in invasive ductal carcinoma (35). Cathepsin D, a protease suggested to have a role in breast cancer, appears to affect the potential for invasive growth (11, 14, 36). Positive immunostaining of tumor sections with Factor VIII antibodies seems to be a marker for angiogenesis (37–39). It has been postulated that these tumors are targets for antiangiogenesis drug treatment. Expression of the *mdr-1* gene is proposed to be an indicator of multidrug resistance (38–40). Poor response to endocrine therapy has been indicated for urokinase-type plasminogen activator/plasminogen activator inhibitor-1, a plasminogen activator inhibitor (21). Also receiving major attention are the familial breast cancer-related genes *BRCA1* and *BRCA2* (40–42). With the availability of tens of thousands of EST sequences, we have, using differential cDNA sequence, identified a new putative breast cancer marker gene, *BCSG1*, and studied its expression in breast cancer.

The differential cDNA sequencing method described here is a direct approach that utilizes an automatic EST analysis on a pair of cDNA libraries. Unlike previously described methods, the differential cDNA sequencing approach allows one to identify differentially expressed genes or patterns of genes directly from a computer database. With the advancement of more efficient and rapid sequencing technology, the direct differential cDNA sequencing approach may offer a powerful method for simultaneous analysis of the expression profile of thousands of genes, as well as for the discovery of novel genes of clinical interest.

Using *in situ* hybridization analysis, we have demonstrated the expression of BCSG1 transcripts in the neoplastic epithelial cells of infiltrating breast carcinoma but not in epithelial cells of normal and benign breast. The overexpression (17 of 20) of BCSG1 in malignant infiltrating breast epithelial cells compared to the partial expression (7 of 15) in *in situ* carcinoma suggests that up-regulation of BCSG1 expression is associated with breast cancer malignant progression and may signal the more advanced invasive/metastatic phenotype of human breast cancer. This implication is supported further by the detection of BCSG1 expression in six of eight aggressive Comedo-type DCISs and in only one of seven non-Comedo type DCISs. It is unlikely that BCSG1 is overexpressed as a secondary effect of cellular proliferation, because no detectable BCSG1 expression is evident in rapidly proliferating nonmalignant breast lesions (Fig. 4*F*).

It will be interesting to investigate whether BCSG1 expression in DCIS may indicate a malignant progression leading to invasion and metastasis. There is cause for concern about the large number of DCIS cases that are being diagnosed as a consequence of screening mammography, most of which are treated by some form of surgery. In addition, the proportion of cases treated by mastectomy may be inappropriately high (30). DCIS by definition has intact basement membrane by light microscopy (43). Defective basement membranes, however, have been found when they are stained with periodic acid-Schiff reagent and when they are examined by electron microscopy (44). In fact, it has been reported that re-evaluation by experienced pathologists showed that 28 and 15% of previously diagnosed DCISs demonstrated invasion (45, 46). If BCSG1 expression can provide some prognostic information on distinguishing the DCIS that is not likely to become invasive from the DCIS that is most likely to become invasive, this will help to direct the treatment strategies and to reduce some inappropriate or unnecessary mastectomies.

It is interesting to note that the predicted amino acid sequence of *BCSG1* gene shares a high sequence homology with the non-Aβ component of the AD amyloid precursor protein (28). A neuropathological hallmark of AD is a widespread amyloid deposition resulting from β-amyloid precursor proteins. β-Amyloid precursor proteins are large, membrane-spanning proteins that either give rise to the β-A4 peptide (Aβ fragment; Ref. 47) or a non-Aβ component of AD amyloid (28) that is either deposited in AD amyloid plaques or yielding soluble forms. Although the insoluble membrane-bound AD amyloid destabilizes calcium homeostasis and thus renders cell vulnerable to excitotoxic conditions of calcium influx resulting from energy deprivation or overexcitation (48), the soluble AD amyloid proteins are neuroprotective against glucose deprivation and glutamate toxicity, perhaps through their ability to lower the intraneuronal calcium concentration (49). We currently do not know whether BCSG1 is an instigator or a by-product during breast cancer progression. With the availability of anti-BCSG1 antibody to localize BCSG1 protein and the recombinant BCSG1 protein, we may start to speculate that BCSG1, like soluble AD amyloid, may be potentially involved in protection from tissue damage resulting from tissue remodeling due to the local cancer invasion. An elucidation of the reasons for BCSG1 overexpression in infiltrating breast cancer cells may shed some light on the pathogenesis of breast cancer progression. Nevertheless, we demonstrated a stage-specific BCSG1 expression and an association of BCSG1 overexpression with clinical aggressiveness of breast cancers. The notion that the BCSG1 overexpression may indicate breast cancer malignant progression from benign breast or *in situ* carcinoma to the highly infiltrating carcinoma warrants further investigation.

ACKNOWLEDGMENTS

We thank Drs. W. Cance and E. Liu (University of North Carolina at Chapel Hill) for providing normal breast tissue and breast cancer for constructing cDNA libraries, Dr. G. Yu and Dr. L. Xing for technical assistance, and Jane Shirreffs for reading the manuscript. The sequence analysis was conducted at the DNA Sequencing Core Facility at Human Genome Sciences, Inc., and The Institute of Genomic Research.

REFERENCES

1. Porter-Jordan, K., and Lippman, M. E. Overview of the biological markers of breast cancer. Hematol. Oncol. Clin. North Am., *8*: 73–100, 1994.
2. Manning, D. L., McClelland, R. A., Knowlden, J. M., Bryant, S., Green, C. D., Sutherland, R. L., and Ormandy, C. J. Differential expression of oestrogen regulated genes in breast cancer. Acta Oncol., *34*: 641–646, 1995.
3. Watson, M. A., and Fleming, T. P. Isolation of differentially expressed sequence tags

from human breast cancer. Cancer Res., *54:* 4598–4602, 1994.
4. Sager, R., Anisowicz, A., Neveu, M., Liang, P., and Sotiropoulou, G. Identification by differential display of α6 integrin as a candidate tumor suppressor gene. FASEB J., *7:* 964–970, 1993.
5. Chen, Z., and Sager, R. Differential expression of human tissue factor in normal mammary epithelial cells and in carcinomas. Mol. Med., *1:* 153–160, 1995.
6. Zhang, M., Zou, Z., Maass, N., and Sager, R. Differential expression of elafin in human normal mammary epithelial cells and carcinomas is regulated at the transcriptional level. Cancer Res., *55:* 2537–2541, 1995.
7. Zou, Z., Anisowicz, A., Hendrix, M. J., Thor, A., Neveu, M., Sheng, S., Rafidi, K., Seftor, E., and Sager, R. Maspin, a serpin with tumor-suppressing activity in human mammary epithelial cells. Science (Washington DC), *263:* 526–529.
8. Liang, P., Averboukh, L., Keyomarsi, K., Sager, R., and Pardee, A. B. Differential display and cloning of messenger RNAs from human breast cancer *versus* mammary epithelial cells. Cancer Res., *52:* 6966–6968, 1992.
9. Adams, M. D., Soares, M. B., Kerlavage, A. R., Field, C., and Venter, J. C. Rapid cDNA sequencing (expressed sequence tags) from a directionally cloned human infant brain cDNA library. Nat. Genet., *4:* 373–380, 1993.
10. Adams, M. D., Kelley, J. M., Gocayne, J. D., Dubnick, M., Polymeropoulos, M. H., Xiao, H., Merril, C. R., Wu, A., Olde, B., Moreno, R. F., Kerlavage, A. R., McCombie, W. R., and Venter, J. C. Complementary DNA sequencing: expressed sequence tags and human genome project. Science (Washington DC), *252:* 1651–1656, 1991.
11. Velculescu, V. E., Zhang, L., Vogelstein, B., and Kinzler, K. W. Serial analysis of gene expression. Science (Washington DC), *270:* 484–487, 1995.
12. Schena, M., Shalon, D., Davis, R. W., and Brown, P. O. Quantitative monitoring of gene expression patterns with a complementary DNA microarray. Science (Washington DC), *270:* 467–470, 1995.
13. Chomczynski, P., and Sacchi, N. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Anal. Biochem., *162:* 156–159, 1987.
14. Angerer, M. L., and Angerer, R. C. *In situ* hybridization to cellular RNA. *In:* D. Rickwood and B. D. Hames (eds.), *In Situ* Hybridization, pp. 15–32. London: LRL Press, 1992.
15. Rochefort, H., Capony, F., Garcia, M., Cavailles, V., Freiss, G., Chambon, M., Morisset, M., and Vignon, F. Estrogen-induced lysosomal proteases secreted by breast cancer cells: a role in carcinogenesis? J. Cell. Biochem., *35:* 17–29, 1987.
16. Cavailles, V., Augereau, P., and Rochefort, H. Cathepsin D gene of human MCF-7 cells contains estrogen-responsive sequences on its 5′ proximal flanking region. Biochem. Biophys. Res. Commun., *174:* 816–824, 1991.
17. Capony, F., Rougeot, C., Cavailles, V., and Rochefort, F. Estradiol increases the secretion by MCF-7 cells of several lysosomal pro-enzymes. Biochem. Biophys. Res. Commun., *171:* 972–978, 1990.
18. Brouillet, J. P., Theillet, C., Maudelonde, T., Defrenne, A., Simony-Lafontaine, J., Sertour, J., Pujol, H., Jeanteur, P., and Rochefort, H. Cathepsin D assay in primary breast cancer and lymph nodes: relationship with c-myc, *c-erb*-B-2 oncogene amplification and node invasiveness. Eur. J. Cancer, *26:* 437–441, 1990.
19. Spyratos, F., Maudelonde, T., Brouillet, J. P., Brunet, M., Defrenne, A., Andrieu, C., Hacene, K., Desplaces, A., Rouesse, J., and Rochefort, H. Cathepsin D, an independent prognostic factor for metastasis of breast cancer. Lancet, *II:* 1115–1118, 1989.
20. Rochefort, H., Cavailles, V., Augereau, P., Capony, F., Maudelonde, T., Touitou, I., and Garcia, M. Overexpression and hormonal regulation of pro-cathepsin D in mammary and endometrial cancer. J. Steroid Biochem., *34:* 177–182, 1989.
21. Foekens, J. A., Look, M. P., Peters, H. A., Van Putten, W. L., Portengen, H., and Klijn, J. G. Urokinase-type plasminogen activator and its inhibitor PAI-1: predictor of poor response to tamoxifen therapy in recurrent breast cancer. J. Natl. Cancer Inst., *87:* 751–756, 1995.
22. Horan-Hand, P., Colcher, D., Salomon, D., Ridge, J., Noguchi, P., and Schlom, J. Influence of spatial configuration of carcinoma cell populations on the expression of a tumor-associated glycoprotein. Cancer Res., *45:* 833–840, 1986.
23. Hunt, G. The role of laminin in cancer invasion and metastasis. Exp. Cell Biol., *57:* 165–176, 1989.
24. Castronovo, V., Colin, C., Claysmith, A. P., Chen, P. H., Lifrange, E., Lambotte, R., Krutzsch, H., Liotta, L. A., and Sobel, M. E. Immunodetection of the metastasis-associated laminin receptor in human breast cancer cells obtained by fine-needle aspiration biopsy. Am. J. Pathol., *137:* 1373–1381, 1990.
25. Marques, L. A., Franco, E. L., Torloni, H., Brentani, M. M., da Silva-Neto, J. B., and Brentani, R. R. Independent prognostic value of laminin receptor expression in breast cancer survival. Cancer Res., *50:* 1479–1483, 1990.
26. Gasparini, G., Barbareschi, M., Boracchi, P., Bevilacqua, P., Verderio, P., Dalla-Palma, P., and Menard, S. 67 kDa laminin-receptor expression adds prognostic information to intra-tumoral microvessel density in node-negative breast cancer. Int. J. Cancer., *60:* 604–610, 1995.
27. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. Basic local alignment search tool. J. Mol. Biol., *215:* 403–410, 1990.
28. Ueda, K., Fukushima, H., Masliah, E., Xia, Y., Iwai, A., Yoshimoto, M., Otero, D. A., Kondo, J., Ihara, Y., and Saitoh, T. Molecular cloning of cDNA encoding an unrecognized component of amyloid in Alzheimer disease. Proc. Natl. Acad. Sci. USA, *90:* 11282–11286, 1993.
29. Goldhirsch, A. Meeting highlights (St. Gallen 1995): International Consensus Panel on the Treatment of Primary Breast Cancer (Commentary). J. Natl. Cancer Inst., *97:* 1141–1145, 1995.
30. Ernster, V. L., Barclay, J., Kerlikowske, K., Grady, D., and Henderson, C. Incidence of and treatment for ductal carcinoma *in situ* of the breast. J. Am. Med. Assoc., *275:* 913–918, 1996.
31. Allred, D. C., Clark, G. M., Tandon, A. K., Molina, R., Tormey, D. C., Osborne, C. K., Gilchrist, K. W., Mansour, E. G., Abeloff, M., and Eudey, L. Her-2/neu in node-negative breast cancer: prognostic significance of overexpression influenced by the presence of *in situ* carcinoma. J. Clin. Oncol., *10:* 599–605, 1992.
32. Gusterson, B. A., Gelber, R. D., Goldhirsch, A., Price, K. N., Soderborgh, J., Anbazhagan, R., Styles, J., Rudenstam, C. M., Golouh, R., and Reed, R. Prognostic importance of c-erbB-2 expression in breast cancer. J. Clin. Oncol., *10:* 1049–1056, 1992.
33. Bergh, J., Norberg, T., Sjogren, S., Lindgren, A., and Holmberg, L. Complete sequencing of the *p53* gene provides prognostic information in breast cancer patients, particularly in relation to adjuvant systemic therapy and radiotherapy. Nat. Med., *10:* 1029–1034, 1995.
34. Elledge, R. M., Fuqua, S. A., Clark, G. M., Pujol, P., and Allred, D. C. The role and prognostic significance of *p53* gene alterations in breast cancer. Breast Cancer Res. Treat., *27:* 95–102, 1993.
35. Steeg, P. S., De La Rosa, A., Flatow, U., MacDonald, N. J., Benedict M, and Leone A. Nm23 and breast cancer metastasis. Breast Cancer Res. Treat., *25:* 175–187, 1993.
36. Ferno, M., Baldetorp, B., Borg, A., Brouillet, J. P., Olsson, H., Rochefort, H., Sellberg, G., Sigurdsson, H., and Killander, D. Cathepsin D, both a prognostic factor and a predictive factor for the effect of adjuvant tamoxifen in breast cancer: South Sweden Breast Cancer Group. Eur. J. Cancer, *30A:* 2042–2048, 1994.
37. Klijn, J. G. M., Berns, E. M. J. J., and Foekens, J. A. Prognostic factors and response to therapy in breast cancer. Cancer Surv. *18:* 165–198, 1993.
38. Harris, A. L., Fox, S., Leek, R., Zhang, H., Scott, P., Bicknell, R., and Gatter, K. Breast cancer angiogenesis: therapy target and prognostic factor. Eur. J. Cancer, *31A:* 831–832, 1995.
39. Gasparini, G., Pozza, F., and Harriss, A. L. Evaluating the potential usefulness of new prognostic and predictive indicators in node-negative breast cancer patients. J. Natl. Cancer Inst., *85:* 1206–1219, 1993. (Erratum, J. Natl. Cancer Inst., *85:* 1605, 1993).
40. Miki, Y., Swenson, J., Shattuck-Eidens, D., Futreal, P. A., Harshman, K., Tavtigian, S., Liu, Q., Cochran, C., Bennett, L. M., Ding, W., Bell, R., Rosenthal, J., Hussey, C., Tran, T., McClure, M., Frye, C., Hattier, T., Phelps, R., Haugen-Strano, A., Katcher, H., Yakumo, K., Gholami, Z., Shaffer, D., Stone, S., Bayer, S., Wray, C., Bogden, R., Dayananth, P., Ward, J., Tonin, P., Narod, S., Bristow, P. K., Norris, F. H., Helvering, L., Morrison, P., Rosteck, P., Lai, M., Barrett, J. C., Lewis, C., Neuhausen, S., Cannon-Albright, L., Goldgar, D., Wiseman, R., Kamb, A., and Skolnick, M. H. A strong candidate for the breast and ovarian cancer susceptibility gene *BRCA1*. Science (Washington DC), *266:* 66–71, 1994.
41. Wooster, R., Neuhausen, S. L., Mangion, J., Quirk, Y., Ford, D., Collins, N., Nguyen, K., Seal, S., Tran, T., and Averill, D. Localization of a breast cancer susceptibility gene, *BRCA2*, to chromosome 13q12–13. Science (Washington DC), *265:* 2088–2090, 1994.
42. Futreal, P. A., Liu, Q., and Shattuck-Eidens, D. BRCA1 mutations in primary breast and ovarian carcinomas. Science (Washington DC), *266:* 120–122, 1994.
43. Ozzello, L. Intraepithelial carcinoma of the breast. *In:* K. H. Hollman and J. M. Verley (eds.), New Frontiers in Mammary Pathology. New York: Plenum Publishing Corp., 1983.
44. Barsky, S. H., Siegal, G. P., Jannotta, F., and Liotta, L. A. Loss of basement membrane components by invasive tumors but not by their benign counterparts. Lab. Invest., *49:* 140–147, 1983.
45. Brown, P. W., Silverman, J., Owens, E., Tabor, D. C., Terz, J. J., and Lawrence, W., Jr. Intraductal "non-filtrating" carcinoma of the breast. Arch. Surg., *111:* 1063–1067, 1976.
46. Lagios, M. D. Carcinoma *in situ:* pathology and treatment. Surg. Clin. North Am., *70:* 853–871, 1990.
47. Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. L., and Beyreuther, K. Amyloid plaque core protein in Alzheimer disease and Down syndrome. Proc. Natl. Acad. Sci. USA, *82:* 4245–4249, 1985.
48. Mattson, M. P., Rydel, R. E., Lieberburg, I., and Smith-Swintosky, V. L. Altered calcium signaling and neuronal injury: stroke and Alzheimer's disease as example. Ann. NY Acad. Sci., *679:* 1–21, 1993.
49. Barger, S. W., Fiscus, R. P., Ruth, P., Hofmann, F., and Mattson, M. P. Role of cyclic GMP in the regulation of neuronal calcium and survival by secreted forms of β-amyloid precursor. J. Neurochem., *64:* 2087–2096, 1995.

BIOCHEMISTRY

SECOND EDITION

CHRISTOPHER K. MATHEWS
OREGON STATE UNIVERSITY

K. E. VAN HOLDE
OREGON STATE UNIVERSITY

THE BENJAMIN/CUMMINGS PUBLISHING COMPANY, INC.
Menlo Park, California • Reading, Massachusetts • New York • Don Mills, Ontario
Wokingham, U.K. • Amsterdam • Bonn • Paris • Milan • Madrid • Sydney
Singapore • Tokyo • Seoul • Taipei • Mexico City • San Juan, Puerto Rico

Sponsoring Editor: Anne Scanlan-Rohrer
Senior Developmental Editor: Shelley Parlante
Associate Editor: Leslie With
Editorial Assistants: Clary Alward, Karmen Butterer, Sharon Sforza
Production Editor: Donna Linden
Art Editor: Kelly Murphy
Photo Editor: Kathleen Cameron
Permissions Editor: Marty Granahan
Composition and Film Manager: Lillian Hom
Composition and Film Assistant: Vivian McDougal
Manufacturing Supervisor: Merry Free Osborn
Marketing Manager: Nathalie Mainland
Developmental Editors: Robin Fox, Margot Otway
Developmental Artist: Darwen Hennings
Copy Editor: Rosemary Sheffield
Proofreaders: Margaret Moore, Alan Titche
Indexer: Shane-Armstrong Information Systems
Text Designer: Bruce Kortebein, Design Office
Illustrations: Precision Graphics, Thompson Type
Page Layout: Curt Boyer
Composition: Thompson Type
Film: H&S Graphics
Printing: von Hoffman
Cover Designer: Yvo Riezebos
Cover Image: Structure of the β subunit of *E. coli* DNA polymerase III holoenzyme, as revealed by x-ray crystallography
Cover Photograph: Courtesy of J. Kuriyan, *Cell* (1992) 69:425–437

Library of Congress Cataloging-in-Publication Data
Mathews, Christopher K., 1937–
Biochemistry / Christopher K. Mathews, K. E. van Holde. — 2nd ed.
p. cm.
Includes bibliographical references and index.
ISBN 0-8053-3931-0
1. Biochemistry. I. van Holde, K. E. (Kensal Edward), 1928– .
II. Title.
QP514.2.M384 1995
574.19'2—dc20 95–18381

3 4 5 6 7 8 9 10—VH—99 98 97

The Benjamin/Cummings Publishing Company, Inc.
2725 Sand Hill Road
Menlo Park, CA 94025

CHAPTER 6

THE THREE-DIMENSIONAL STRUCTURE OF PROTEINS

In Chapter 5 we introduced the concept of protein primary structure. We emphasized that this first level of organization, the amino acid sequence, is dictated by the DNA sequence in the gene for each protein. Most proteins exhibit higher levels of structural organization. It is the specific three-dimensional structure of each protein that allows it to function in its particular biological role.

Figure 6.1 depicts another representation of the three-dimensional conformation of the myoglobin molecule we showed in Figure 5.1. A specific 3-D structure means that every one of the thousands of atoms in a protein molecule has a particular, well-defined spatial location within the molecule. This characteristic is emphasized in Figure 5.1. Figure 6.1, on the other hand, has been drawn to point out that there exist two distinguishable levels of three-dimensional folding of the polypeptide chain. First, the chain appears to be locally coiled into regions of helical structure (labeled A–H in the figure). Such local *regular* folding is called the secondary structure of the molecule. The helically coiled regions are in turn folded into a specific compact structure for the entire polypeptide chain. We call this further level of folding the tertiary structure of the molecule. Later in this chapter we shall find that some proteins consist of several polypeptide chains, arranged in a regular manner. This arrangement we designate as the quaternary level of organization.

This chapter is devoted to an examination of the several levels of protein structure—their geometry, how they are stabilized, and their importance in protein function.

Protein molecules have four levels of structural organization: primary (sequence), secondary (local folding), tertiary (overall folding), and quaternary (multichain association).

SECONDARY STRUCTURE: REGULAR WAYS TO FOLD THE POLYPEPTIDE CHAIN

THE DISCOVERY OF REGULAR POLYPEPTIDE STRUCTURES

Our understanding of the protein secondary structure had its origins in the remarkable work of Linus Pauling, perhaps the greatest chemist of the twentieth

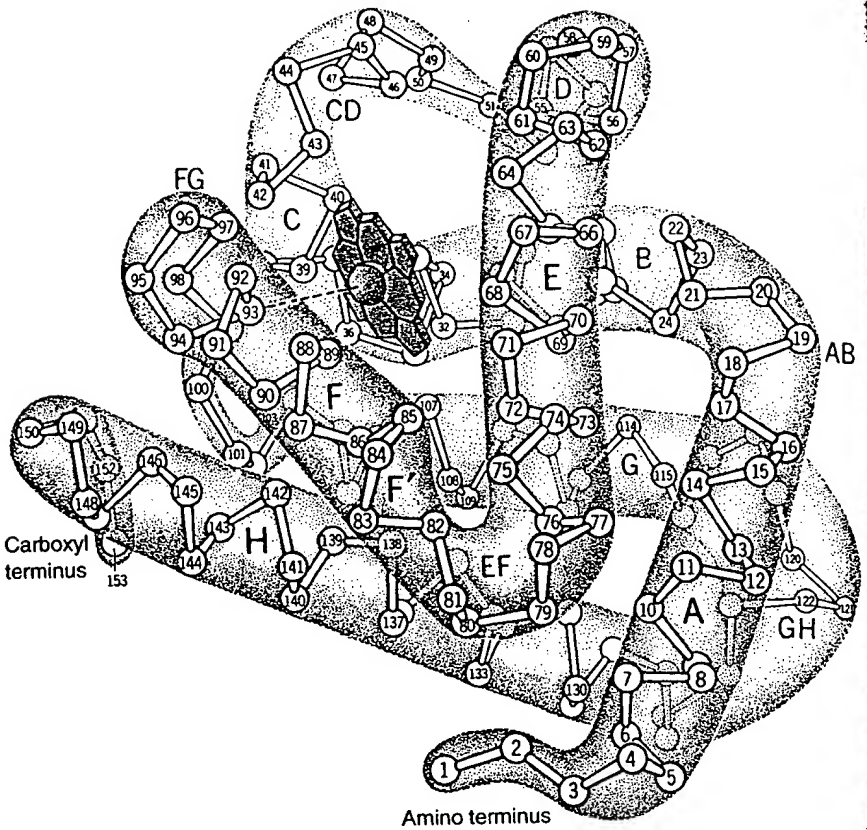


FIGURE **6.1**

Three-dimensional folding of the protein myoglobin. Each amino acid is indicated by a circle corresponding to its α carbon atom. Side chains are omitted, to emphasize how the polypeptide backbone is wrapped into helices and folded. Individual α-helical regions are labeled A–H, with turn regions designated by two letters (e.g., GH). This protein folds about a heme group (shown in purple), a planar heterocyclic structure that chelates iron and serves as the oxygen binding site.

century. As early as the 1930s, he had begun x-ray diffraction studies of amino acids and small peptides, with the aim of eventually analyzing protein structure. In the early 1950s, Pauling and his collaborators used these data together with unusual scientific intuition to begin a systematic analysis of the possible regular conformations of the polypeptide chain. They postulated several principles that any such structure must obey:

1. The bond lengths and bond angles should be distorted as little as possible from those found through x-ray diffraction studies of amino acids and peptides, as shown in Figure 5.12b (page 141).
2. No two atoms should approach one another more closely than is allowed by their van der Waals radii.
3. The amide group must remain planar and in the trans configuration, as shown in Figure 5.12b. (This feature had been recognized in the earlier x-ray diffraction studies of small peptides.) Consequently, rotation is possible only about the two bonds adjacent to the α carbon in each amino acid residue, as shown in Figure 6.2.

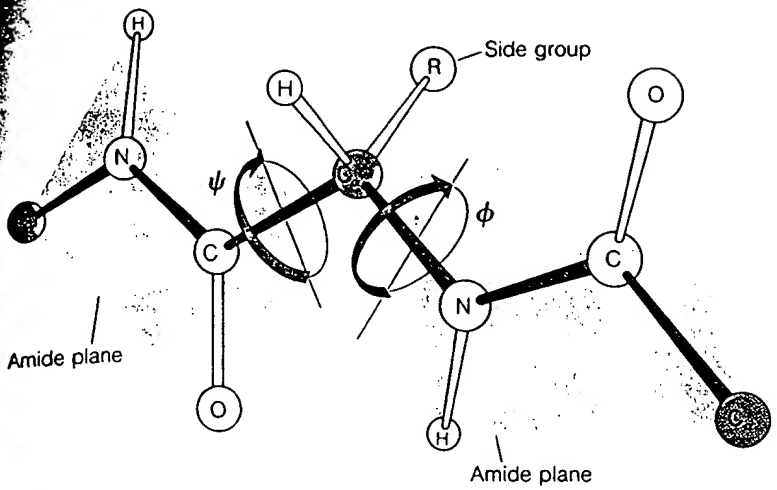
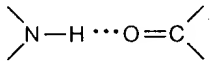


FIGURE 6.2

Rotation around the bonds in a polypeptide chain. Two adjacent amide planes are shown in teal. Rotation is allowed only about the N—C_α and C_α—C(=O)— bonds. The angles of rotation about these bonds are defined as ϕ and ψ, respectively, with directions defined as positive rotation as shown by the arrows. The extended conformation of the chain shown here corresponds to $\phi = +180°$, $\psi = +180°$.

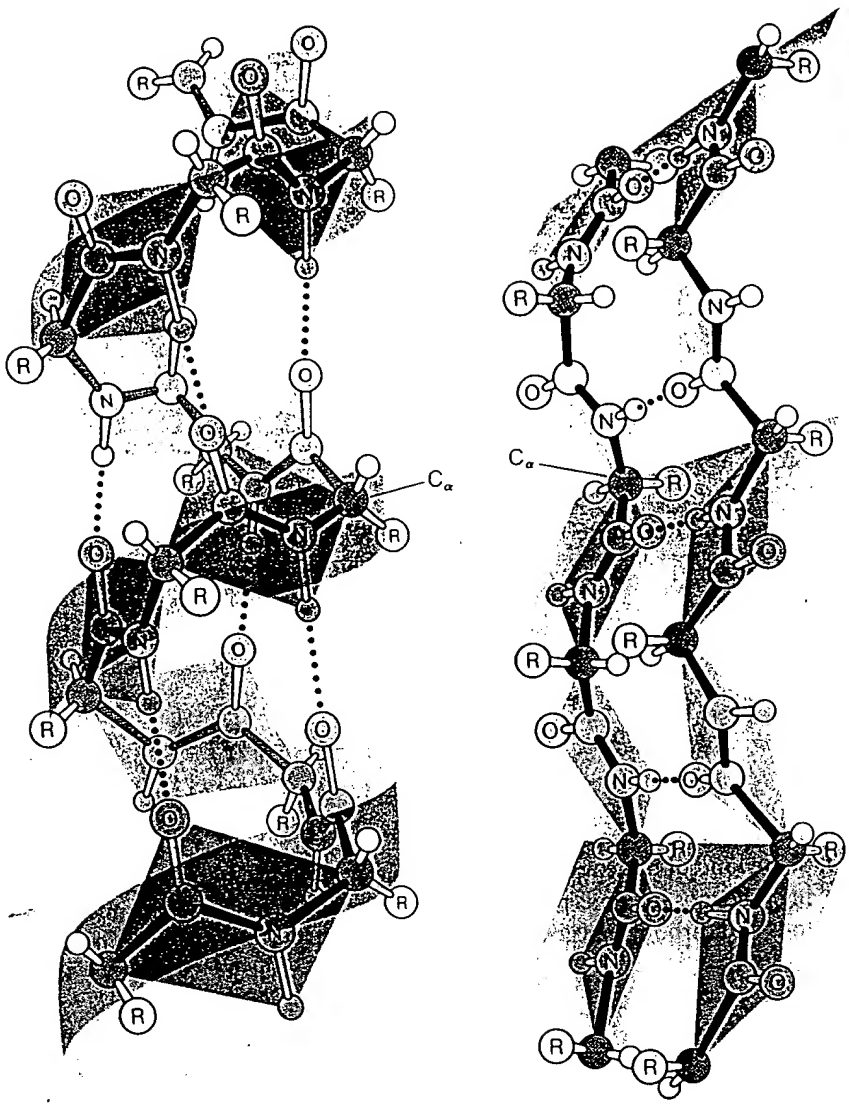
4. Some kind of noncovalent bonding is necessary to stabilize a regular folding. The most obvious possibility is hydrogen bonding between amide protons and carbonyl oxygens:



Such a concept was natural to Pauling, who had had much to do with the development of the idea of H bonds. In summary, the preferred conformations must be those that allow a maximum amount of hydrogen bonding, yet satisfy criteria 1–3.

Working mainly with molecular models, Pauling and his associates were able to arrive at a small number of regular conformations that satisfied all of these criteria. Some were helical structures formed by a single polypeptide chain, and some were sheetlike structures formed by adjacent chains. The two structures they proposed as most likely—the *α helix* and *β sheet*—are shown in Figure 6.3a and b. These two structures turned out, in fact, to be the most common secondary structures in proteins. Figure 6.4 shows two other polypeptide helices that have since been defined. The 3_{10} helix is observed in some proteins but is not as common as the α helix. The π helix, though sterically possible, has not been observed, possibly because it has a hole down the middle too big to allow van der Waals interactions but too small to admit potentially stabilizing water molecules. All of the structures shown in Figures 6.3 and 6.4 satisfy the criteria listed earlier. In particular, in each structure the peptide group is planar, and every amide proton and every carbonyl oxygen (except a few near the ends of helices) is involved in hydrogen bonding. Each of these forms constitutes a possible kind of secondary structure in proteins.

Of the several possible secondary structures for polypeptides, the most important are the α helix, the β sheet, and the 3_{10} helix.



(a) α helix (b) β sheet

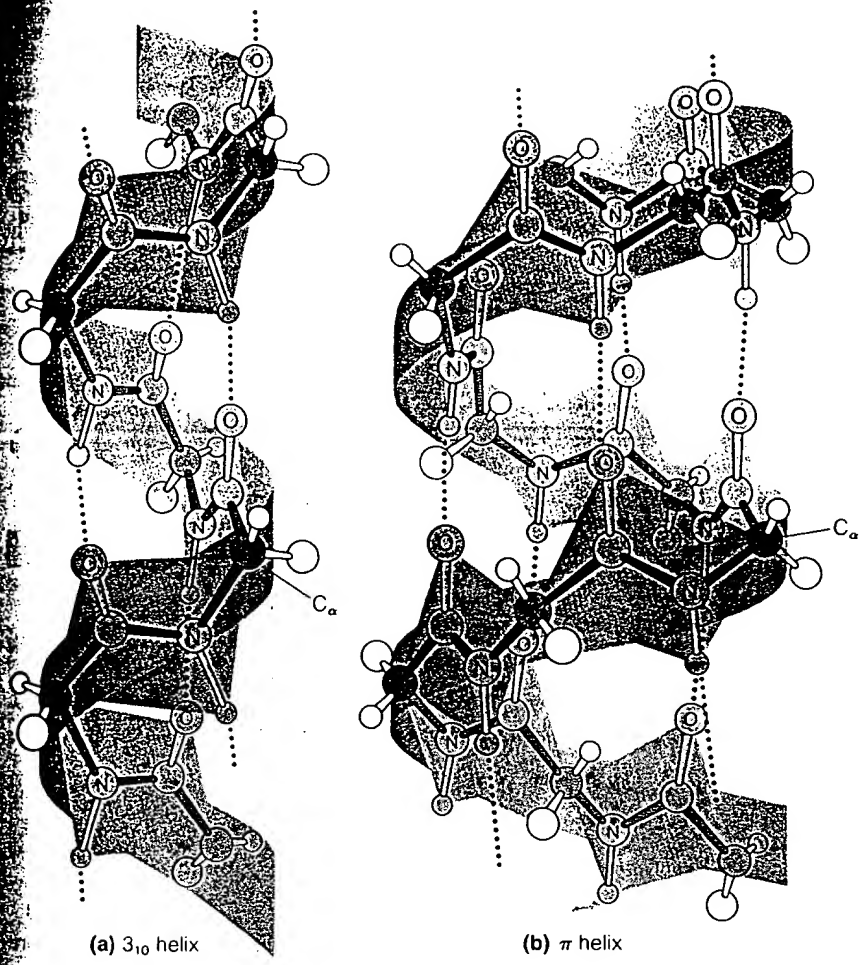
FIGURE 6.3

The α helix and β sheet. These are the two most important regular secondary structures of polypeptides. Their existence was predicted by Linus Pauling. **(a)** In the α helix the hydrogen bonds are within a single polypeptide chain. **(b)** In the β sheet, the hydrogen bonds are between adjacent chains, of which only two are shown here.

DESCRIBING THE STRUCTURES: MOLECULAR HELICES AND PLEATED SHEETS

In Tools of Biochemistry 4B, we listed the distances that define a molecular helix the crystallographic repeat (c), the pitch (p), and the rise (h). We also pointed ou that helices may be either right-handed or left-handed and may contain either a integral number of residues per turn or a nonintegral number. We call the num ber of residues per turn n and the number of residues per repeat m. The latt number must always be an integer, because it defines an exact repeat of the stru ture. If there is an integral number of residues per turn, the pitch and the repe will be equal, and $n = m$. Some helices of this kind are illustrated schematically Figure 6.5. Note that as the number of residues per turn increases, the structu changes progressively from a flat ribbon to a broad helix and eventually to closed ring with $p = 0$. Not all of these structures are found in polypeptides. F example, the single-chain $n = 2$ structure shown in Figure 6.5 does not occur nature, as we will explain shortly.

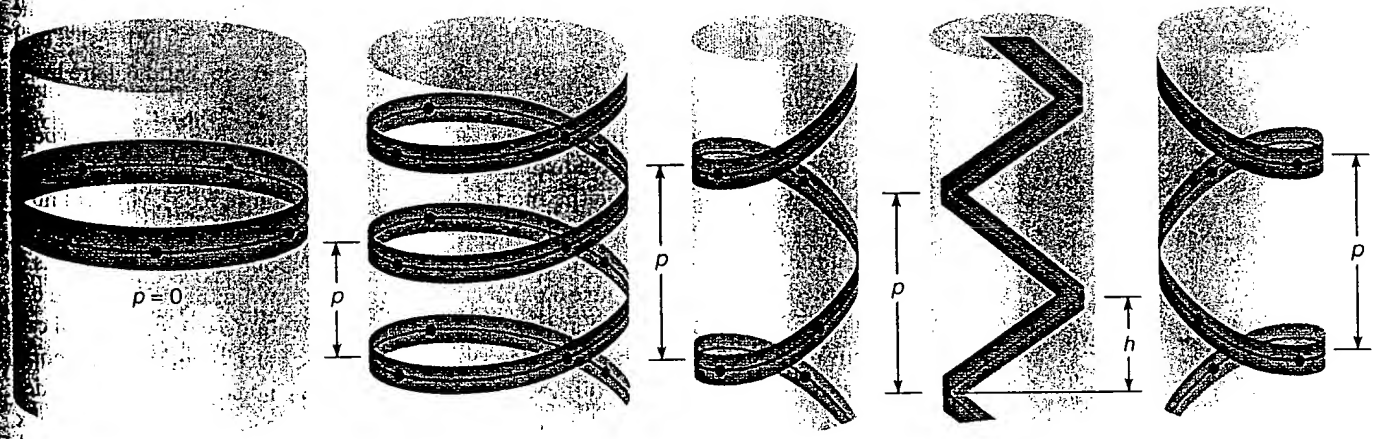
One of Pauling's major insights was to recognize that polypeptide helices not have to have an integral number of residues per turn. For example, the



(a) 3_{10} helix (b) π helix

FIGURE **6.4**

Other possible secondary structures of polypeptides. **(a)** The 3_{10} helix is found in proteins but is less common than the α helix shown in Figure 6.3a. **(b)** The π (or 4.4_{16}) helix is sterically possible but so far has not been observed in proteins.



$n = 5$ Ring | $n = 4$ Helix (right-handed) | $n = 3$ Helix (right-handed) | $n = 2$ Flat ribbon | $n = -3$ Helix (left-handed)

FIGURE **6.5**

Idealized helices. These hypothetical structures show the effect of varying the number (n) of polypeptide residues per turn of a helix. In each case the pitch (p) is indicated, and for $n = 2$ the rise (h) is also shown. Polypeptides can form helices ranging from a closed ring ($n = 5$, $p = 0$) to a 2-fold helix ($n = 2$), with each residue rotated by 180° with respect to the preceding one. All the integral positive values of n and one example of a negative value are shown here. The $n = 4$ and $n = 3$ helices are right-handed, the $n = -3$ helix is left-handed, and $n = 5$ (a ring) and $n = 2$ (a ribbon) have no handedness. The right-handed α helix (not shown here), with $n = 3.6$, is intermediate between the $n = 3$ and $n = 4$ structures.

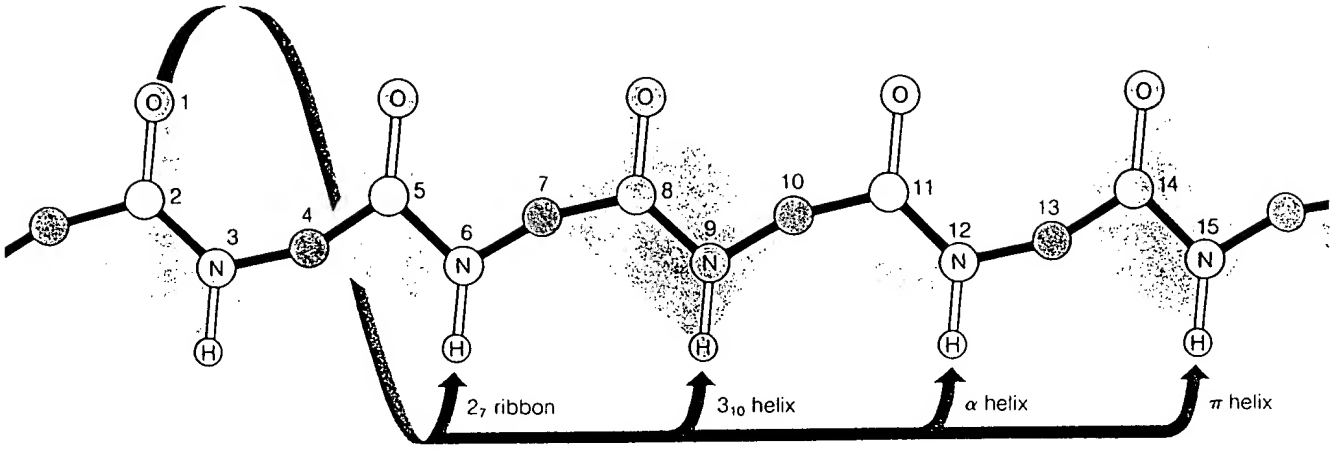


FIGURE **6.6**

Hydrogen bonding patterns for four helices. The structures are represented in a diagrammatic way to simplify counting the atoms in each H-bonded loop. For example, there are 13 atoms in the H loop corresponding to the α (3.6_{13}) helix.

TABLE 6.1 Parameters of some polypeptide secondary structures

Structure Type	Residues/ Turn	Rise (nm)	Number of Atoms in H-Bonded Ring	ϕ (°)	ψ (°)
Antiparallel β sheet	2.0	0.34	—[a]	−139	+135
Parallel β sheet	2.0	0.32	—[a]	−119	+113
3_{10} helix	3.0	0.20	10	−49	−26
α helix (3.6_{13})	3.6	0.15	13	−57	−47
π helix (4.4_{16})[b]	4.4	0.12	16	−57	−70

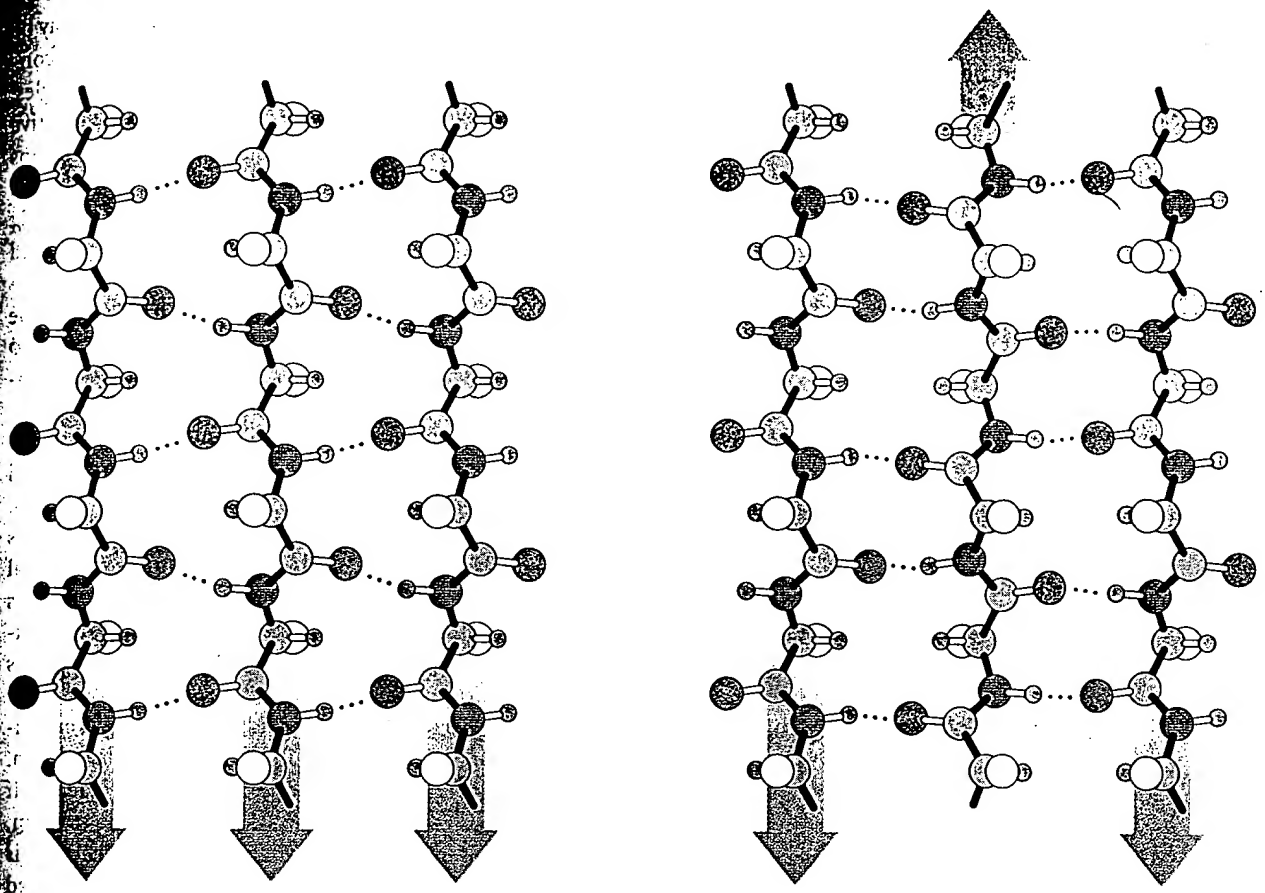
[a]Bonding is between polypeptide chains.

[b]Sterically permitted but not observed in protein.

helix repeats after exactly 18 residues, which amounts to 5 turns. It has, therefore, 3.6 residues per turn. Since the pitch of a helix is given by $p = nh$, we have for the α helix, with a rise of 0.15 nm/residue, p = 3.6 (res/turn) × 0.15 (nm/res) = 0.54 nm/turn. Parameters for the other helices shown in Figures 6.3 and 6.4 are listed in Table 6.1.

The parameters defined above describe most molecular helices. For polypeptide helices, which involve hydrogen bonding, there is an additional important quantity. If you examine the model for the α helix (Figure 6.3a), you will note that each carbonyl oxygen is hydrogen-bonded to the amido proton on the *fourth* residue up the helix. Thus, if we include the hydrogen bond, a loop of 13 atoms is formed, as shown in Figure 6.6. Each of the helices shown in Figures 6.3 and 6.4 has a different number of atoms in such a hydrogen-bonded loop. We shall call this number N. A quick way to describe a polypeptide helix, then, is by the shorthand n_N, where n is the number of residues per turn. The 3_{10} helix fits this description; it has exactly 3.0 residues per turn and a 10-member loop. The α helix could also be called a 3.6_{13} helix, and the π helix a 4.4_{16} helix.

Because hydrogen bonds tend to be linear, the atoms —N—H···O= in polypeptide helices should lie on a straight line. If you examine Figures 6.3 and 6.4, you will see that this requirement is at least approximately satisfied for the 3_{10}, α, and π helices. However, it is very difficult to make helices with only two residues per turn *and* linear hydrogen bonds between residues in the same chain. Therefore, the only $n = 2$ structure that is found in proteins is *not* the flat ribbon shown in Figure 6.5 but the β *pleated sheet* structure shown in Figure 6.3b. In the β pleated sheet, each residue is rotated by 180° with respect to the preceding one, which makes each chain an $n = 2$ helix. If the chains are also folded in an accordion-like



(a) Parallel β sheet (b) Antiparallel β sheet

FIGURE **6.7**

Two kinds of β sheet. The arrows point in the N ⟶ C direction in each chain. Gray = carbon, purple = nitrogen, red = oxygen, blue = hydrogen, white = side chain.

fashion, linear hydrogen bonds can occur *between* adjacent chains. Forming interchain bonds allows correct bond angles with minimal strain when $n = 2$. Figure 6.7 shows the two ways in which this can be done. The chains can have their N ⟶ C directions running parallel, to make the *parallel β sheet*, or they can be *antiparallel*. It is instructive to try to form such structures using molecular models. You will find that making an $n = 2$ helix with internal H bonds is awkward, but the pleated-sheet structures form naturally.

Thus, possible secondary protein structures fall into two general classes: various helices and at least two types of pleated sheet. But just because a structure can be drawn and contains good H bonds does not mean that it necessarily exists. Many kinds of chain conformations can be imagined that are sterically impossible because atoms in the backbone and/or side chains would overlap. These steric restrictions can be fully appreciated only by examining space-filling models, as shown in Figure 6.8. The α helix, for example, is fully packed. To examine steric crowding in a systematic way, we need a general procedure for describing polypeptide conformations.

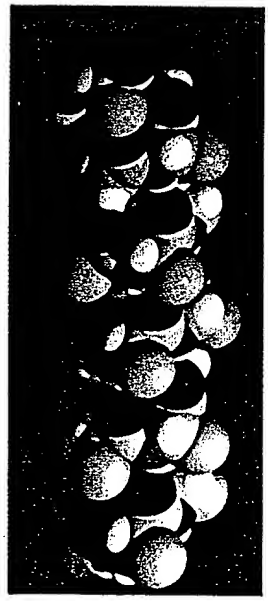


FIGURE **6.8**

Segment of an α helix shown as a space-filling model. The segment illustrated is from the E helix in sperm whale myoglobin (see Figure 6.1).

Courtesy of Richard J. Feldman, National Institutes of Health.

RAMACHANDRAN PLOTS

As was shown in Figure 6.2, each residue in a polypeptide chain has two backbone bonds about which rotation is permitted—the bond between the nitrogen and the α carbon, and the bond between the α carbon and the carbonyl oxygen.